STRUCTURAL STUDIES OF NUCLEOSOMES WITH SMALL ANGLE X-RAY SCATTERING AND CONTRAST VARIATION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Joshua Minoru Tokuda

August 2017

STRUCTURAL STUDIES OF NUCLEOSOMES WITH SMALL ANGLE X-RAY SCATTERING AND CONTRAST VARIATION

Joshua Minoru Tokuda, Ph.D.

Cornell University 2017

Nucleosomes are the fundamental units of DNA packaging and play central roles in regulating genetic activity in eukaryotic cells. High-resolution crystal structures deliver detailed snapshots of the nucleosome in its most stable conformations. However, nucleosomes are dynamic structures that populate diverse conformations as they regulate DNA accessibility. A full understanding of how DNA is processed requires knowledge of these conformations and the interplay of factors that coordinate their formation. In this dissertation, we describe a new approach using small angle x-ray scattering to resolve alternate nucleosome conformations and investigate the mechanisms that facilitate their formation. We begin by studying the salt-induced disassembly pathway of nucleosomes. By modulating the ionic strength of the solution, we globally destabilize nucleosome structures and identify key intermediates that reflect the intrinsic mechanical propensities of the nucleosome. We then investigate how the chromatin remodeler, Chd1, reorganizes nucleosome structure in a nucleotide-dependent manner. Our results suggest that the biophysical properties of the DNA play important roles in directing how alternate conformations are formed.

BIOGRAPHICAL SKETCH

Joshua Minoru Tokuda was raised under the sunny skies of Southern California. From an early age he gained an interest in understanding the natural world and was fortunate enough to be surrounded by friends and family who tolerated his "investigations." His fascination with cell biology began while attending Mr. Seko's frightening AP biology classes at Gahr High School. Despite his grueling training schedule and athletic aspirations to become the next unremarkable runner, he applied to UCLA and entered as an undeclared life sciences major. While attending UCLA, he discovered physics as the ultimate tool for understanding nature and decided to pursue a degree in biophysics. He began his research career in the lab of Dolores Bozovic, where he studied the biophysics of hair cells, the sensory receptors responsible for hearing. After dissecting numerous frog ears and practicing enough Japanese in the lab's acoustic isolation chamber, he received a B.S. in Biophysics and a B.A. in Japanese from UCLA. He began his Biophysics Ph.D. program in 2010 at Cornell University in Ithaca, NY. He joined the lab of Lois Pollack and worked on numerous projects using synchrotron x-rays to study a variety of biological systems. Most importantly, he became very adept at dissolving spoonfuls of sugar and pipetting viscous sucrose solutions. Mary Poppins would be proud.

To My Family

ACKNOWLEDGMENTS

I would like to express my special appreciation and thanks to my advisor Lois Pollack. Your keen guidance, contagious enthusiasm, and relentless compassion has been such an inspiration to me and has deeply shaped who I aspire to be as a person in addition to as a scientist. I would also like to thank all of the Pollack lab members with whom I have been fortunate enough to spend time: Suzette Pabit, Julie Sutton, Yujie Chen, Steve Meisburger, Huimin Chen, Joshua Blose, Chris Jones, Li Li, Andrea Katz, George Calvey, Alex Plumridge, Alex Mauney, Yen Chen, and Josue San Emeterio. You have all made lab such an enjoyable place to work, and I have learned so much from each and every one of you. Special thanks to Yujie Chen, my owl shift partner, who never failed to share my enthusiasm for burning the midnight oil so we can collect as much data as possible until the beam went down.

I thank my special committee members, Brian Crane and Warren Zipfel, for your support, advice, and thoughtful feedback. My first SAXS experiment was in a collaboration with Brian Crane to study the conformational response of cryptochrome. I had the pleasure of working with Anand Vaidya who served as an early student mentor to me. The general strategies we used for cryptochrome inspired the development of structural modeling methods applied in the rest of my work. In my rotation with Warren Zipfel, I not only learned the fundamentals about microscopes and the use of light to study biology, but I also gained an appreciation for instrumentation and the belief that anything can be built better! My time working with you strongly influenced my approach to study biophysics.

My research with nucleosomes would not have been possible with the efforts and patience of my collaborators. For the nucleosome studies in Chapters 3-4, my collaborators from Team NCP West, Lisa Gloss and Traci Topping, prepared the samples, and were responsible for the FRET experiments. For the nucleosome studies in Chapter 5, Greg Bowman and Emily Ren prepared the samples and came to beamtime for the initial experiments. I will never forget Greg's determination in weighing out precise amounts of sucrose on the microgram scale.

I have also benefited from working with brilliant collaborators on numerous projects not included in the main chapters. With Michael Brenowitz and Camille Padlan, I investigated the binding of the RNA aptamer MyrMA with the HIV gag protein. With William Dichtel, Brian Smith, and Ryan Bisbey, we studied the growth of covalent organic frameworks (COF). With Michael Rossmann, Yingyuan "Nick" Sun, we studied DNA ejection and resolved the conformational rearrangements of the capsids from bacteriophage phix174 and ST-1. With Scott Showalter and Eric Gibbs, we investigated the intrinsically disordered C-terminal of RNA polymerase II. I cherished the opportunity to interact with and learn from each and every one of my collaborators.

This work would have not been possible without the support of numerous people who have served as teachers or provided technical, experimental, and administrative assistance. I would like to thank the staff at the Cornell High Energy Synchrotron Source (CHESS), especially Arthur Woll, Richard Gillilan, Ken Finklestein who has taught me more about synchrotrons than I ever expected to know. Thanks to Abdullah Ozer for helping me get started with RNA production by strategizing how to make M-box RNA. Thanks to Nathan Ellis for teaching me how to machine safely and

encouraging my obsessive use of toe clamps. Thanks to Jeff Koski and Mike Cook for machining and welding assistance. I would also like to thank the staff members of the Department of Molecular Biology and Genetics (Vicki Shaff and Ginger Tomassini) and the School of Applied and Engineering Physics (Renee King, Donna Mazza, Bonni Davis, Jon Valezquez, Cynthia Reynolds, Cheri Sigmund, LiLynn Graves), all of whom has supported me on multiple occasions. I am grateful to all of the faculty in the field of biophysics who direct the program and continuously organize opportunities for students to engage in the research community through colloquiums and seminars.

I would like to thank my Ithaca family: Stella Li, Avtar Singh, David Ackerman, Digbijay Majat, Rupa Shah, Jeahoo Kwon, Vinh Le. You have all made Ithaca home for me and I am so grateful and honored to call you friends. You have all been so fun, supportive, and inspiring. I look forward to continuing our adventures!

Finally, I would like to thank my family. My parents have been nothing but supportive, patient, and loving throughout my life. My dad, the hardest worker I know, has been the ultimate provider. He always finds ways to support me without my knowledge, such as getting up extra early in the morning to wipe the condensation off my car windows and secretly fill up my gas tank. My mom is the sweetest woman I know, continuously asking me what I want for the next care package she intends on shipping to me. Undoubtedly, her insistence on making me take swimming lessons as a child has led to *all* of the success in my life. My sister has always been a role model to me and I can always count on her guidance on life. She also saved many of my brain cells by catching me numerous times when I was falling as a toddler, and has rescued me from collapsing cribs and moments of peril on jungle gyms.

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

"In nature's infinite book of secrecy
A little I can read."

William Shakespeare, *Antony and Cleopatra*, 1606

The dazzling beauty of nature inspires a boundless, yet exciting quest to uncover the basis that gives rise to life's diversity. To the delight (and sometimes disdain) of many scientists, this quest has led to an endless cycle of new questions. As newly discovered mechanisms add to an increasingly complex picture of how organisms function, they often underscore our awareness for how much more exists to be discovered. In such a way, research leads many to share the sentiment of the Soothsayer (fortuneteller) quoted above.

Chapter 1

Introduction

In eukaryotic cells, 75-90% of DNA is wrapped in spool-like structures called nucleosomes. As the fundamental organizational unit of DNA, nucleosomes consist of approximately 150 base-pairs of DNA wound around a core of histone proteins. In order to facilitate the vast array of genome-related functions required by the cell, nucleosomes accommodate diverse conformations and compositions. The key feature of nucleosomes is their ability to be modified by extrinsic proteins and host epigenetic marks. This tunable nature provides cells with a powerful platform for manipulating how DNA is accessed and organized in the nucleus. Since the activity of nuclear proteins is influenced by how DNA is packaged, variation of nucleosome structures allows cells to assert control over the expression and heredity of phenotypes beyond what is encoded in the genome.

The organization of DNA in nucleosomes has profound implications for the mechanisms of all cellular activities that require DNA access. Studies using x-ray crystallography have determined high-resolution structures of fully wrapped nucleosomes. These structures reflect the most stable and compact configuration of the nucleosome that serves as the functional form for packaging. Since nucleosomes occlude interactions with the DNA, proteins that catalyze DNA-templated chemistries rely on at least some unwrapping of the DNA from the histones. Research over the past decade has revealed that nucleosomes are highly dynamic structures – a sharp contrast to the static entities they were once thought to be. Although various experimental

approaches report on unique aspects of these dynamics, none of them provide structures of the alternate nucleosome conformations in solution.

The central aim of this work is to gain structural insights into how nucleosomes unwrap to provide DNA access. Specifically, we address two questions:

(1) What are the alternate nucleosome conformations during disassembly?

(2) How do these alternate conformations form?

This chapter lays out the general framework for nucleosome research that motivates our studies. We start by reviewing the nature of DNA and the central role of nucleosomes in the large-scale organization of chromatin. We briefly discuss examples of how nucleosomes regulate gene expression on the level of individual genes. This is followed by a more detailed description of the fully wrapped nucleosome structure and a discussion of how its modular architecture facilitates the formation of alternate configurations. We then summarize studies that explore nucleosome dynamics. We end by outlining the research reported in the following chapters.

1.1 Chromatin and the central role of nucleosomes in DNA compaction

Research over the last century has revealed DNA to be the fundamental carrier of genetic information. Cells have developed remarkable mechanisms to copy, edit, and repair the DNA. As a physical medium for storing information, DNA is remarkably stable (half-life > 500 years) and remains unparalleled in terms of data density. With a diameter of 20 Å and each base-pair contributing a rise of 3.4 Å (10.5 bases per turn) [1], double-stranded DNA has enabled storage of as much as $\approx 5 \times 10^{15}$ bits/mm^3, vastly eclipsing that of commercial flash memory ($\approx 5 \times 10^9$ bits/mm^3) [2]. Nevertheless,

eukaryotic cells require extensive amounts of DNA to store and manage the enormous amount of information required for maintaining homeostasis. In humans, each cell contains nearly 2 linear meters of DNA! Remarkably, this DNA is stored within a cell nucleus with an average diameter of 6 micrometers [3]. This compaction is achieved through the organization of DNA as chromatin – a dynamic arrangement of proteins and nucleic acids within the cell nucleus.

Chromatin is dynamic in both form and function [3]. Its organization adapts to the needs of the cell and facilitates many fundamental cellular processes. For example, during cell division, chromatin compacts into its densest form (metaphase/anaphase chromosomes) to help segregate the chromosome copies into new daughter cells. Since eukaryotic cells spend most of their existence in non-dividing states (resting/interphase), their genomes are generally less condensed than during cell division. Chromatin is kept in either looser ("euchromatin") or tighter ("heterochromatin") structures depending on the metabolic activity level of the genes contained within those regions. Lightly packed chromatin is generally associated with active genes since they permit genetic machinery (such as transcription factors and polymerases) to have regular access to the underlying DNA. One striking illustration of this principle is in cell differentiation. Despite sharing the same genetic code, cells exhibit drastically different phenotypes by manipulating their chromatin architecture on large scales to vary how their genes are expressed [4].

Central to the adaptive nature of chromatin is its hierarchical structural organization (Figure 1.1) [5–7]. In the first level of this hierarchy, approximately 150 base-pairs of

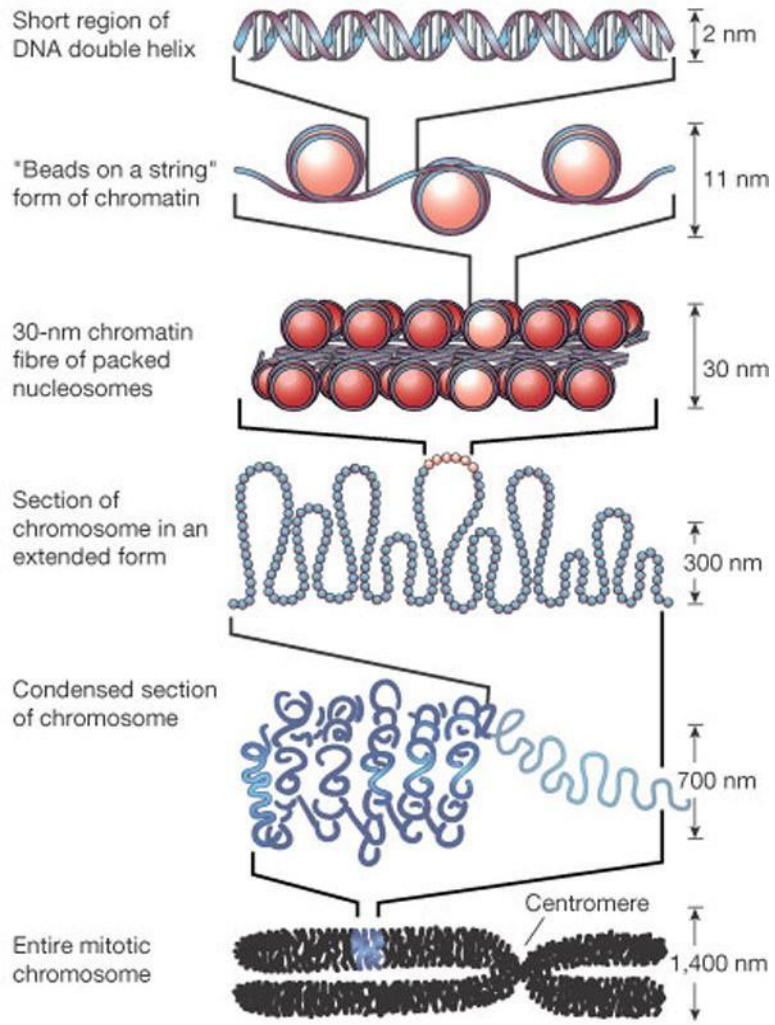


Figure 1.1 The hierarchical organization of eukaryotic chromatin

Adapted by permission from Macmillan Publishers Ltd: Nature [5], © 2003

DNA is wrapped around histone proteins to form spool-like structures called nucleosomes. Tandem nucleosomes are connected by stretches of 10-80 base-pairs of "linker DNA" in a manner that resembles "beads on a string." These nucleosome arrays make up the primary structure of chromatin and are analogous to peptide chains in protein folding. Unlike protein residues, which consist of at least 20 distinct amino acids with diverse biochemical properties, the basic nucleosome-linker repeats are

thought to be largely identical. Recent studies have begun to uncover a host of variable parameters that contribute to nucleosome complexity: DNA sequence [8–10], histone protein variants [11–13], post-translational modifications of histone proteins [14,15], DNA modifications [16,17], and lengths of the linker DNA [18,19]. Through various combinations of these parameters, nucleosomes can be tuned in how they coordinate the formation of functional 3D architectures.

Although full descriptions of higher order structures are yet to be resolved, many in vitro studies have demonstrated the strong propensity of nucleosome arrays to form 30 nm fibers – commonly regarded as the secondary structure of chromatin. Details of this form and its relevance in vivo are active (and controversial) areas of research [20–22]. The difficulty in reaching consensus arises from the intrinsic irregularity of native nucleosomes and the variable lengths of the linker DNA. Of the two most prominent models for 30-nm fiber, the solenoid [23] and zig-zag [24] models, recent studies using highly regular nucleosomes support the latter [25,26]. X-ray diffraction data on crystallized tetranucleosomes (regularly spaced arrays of four nucleosomes) under high ionic strength conditions (90 mM Mg^{2+}) provides the only known high-resolution structure of the fiber and shows the zigzag arrangement [25]. The relevant forms of the fiber with the native heterogeneity expected in long chromatin remain unknown. Beyond the 30 nm fiber, higher-order tertiary structures are formed through long-distance interactions between chromatin fibers [27]. These interactions are facilitated by chromatin associated proteins (CAPs) and scaffold/matrix proteins that further condense chromatin and anchors the DNA onto the nuclear matrix [28–31].

Due to the central role of nucleosomes within this hierarchy, efforts to understand

chromatin architecture depend on accurate models of nucleosomes. Most models of higher order chromatin (including the 30 nm fiber) assume fully wrapped nucleosome structures and do not account for possible variations in their conformations [26]. This assumption stems from the inadequate resolution of individual nucleosomes and linker regions in electron microscopy and x-ray diffraction investigations of fiber structures. One intriguing possibility is that alternate nucleosome configurations may serve as important structural features in chromatin packing. This consideration is largely unexplored, but is potentially important in creating more accurate models of chromatin architecture.

1.1 Nucleosomes regulate genetic activity

In addition to facilitating DNA compaction, nucleosomes play essential roles in regulating genetic activity [32]. This regulation is primarily achieved through its control of DNA access [33–36]. Since histones serve as steric obstacles to enzymes that engage DNA, their strategic placement is a powerful strategy for directing enzymes to appropriate entry sites. In this way, RNA polymerase initiates transcription at the transcription start site (TSS) [37,38], and repair enzymes are guided to sites of DNA damage [39,40]. Genome-wide maps of nucleosomes occupancy and positioning reveal examples of both up- and down-regulation of gene expression [32]. Nucleosomes influence transcription rates by either occluding or directing transcription factors and RNA polymerase to the promoter regions of genes [41,42]. Further control is facilitated through the ability of nucleosomes to recruit proteins, such as pioneer transcription factors [43–46] and to provide sites for epigenetic marks [47–49].

1.2 Nucleosome Structure

High-resolution structures of the "canonical" nucleosome has been solved by x-ray crystallography [36,50–52]. The crystallized structural unit, called the nucleosome core particle (NCP), consists of approximately 146 base-pairs of DNA wrapped in ≈ 1.7 superhelical turns around a core of eight histone proteins. This protein core consists of two copies each of the four histone proteins: H2A, H2B, H3, and H4. As shown in Figure 1.2, NCPs are best characterized by their modular architecture. When formed in vitro, in the absence of interacting factors, the histone proteins first associate into obligate heterodimer pairs of H2A-H2B and H3-H4. Two copies of H3-H4 heterodimers further dimerize to form an $(H3\text{-}H4)_2$ tetramer. The octameric histone core consists of one copy of the $(H3\text{-}H4)_2$ tetramer sandwiched by a copy of an H2A-H2B dimer on both sides. Together, these proteins form a superhelical ramp around which the DNA is tightly wrapped. Nucleosome assembly and re-arrangement is facilitated through the activity of two general classes of proteins: histone chaperones and chromatin remodelers [53]. Histone chaperones bind histone proteins (usually H2A-H2B or H3-H4 heterodimers) and regulate nucleosome assembly by trafficking histones for insertion and preventing aggregation [54–56]. Chromatin remodelers are ATP-driven machines that slide, remodel, and remove nucleosomes, but have also been implicated in chaperone activity [57–60].

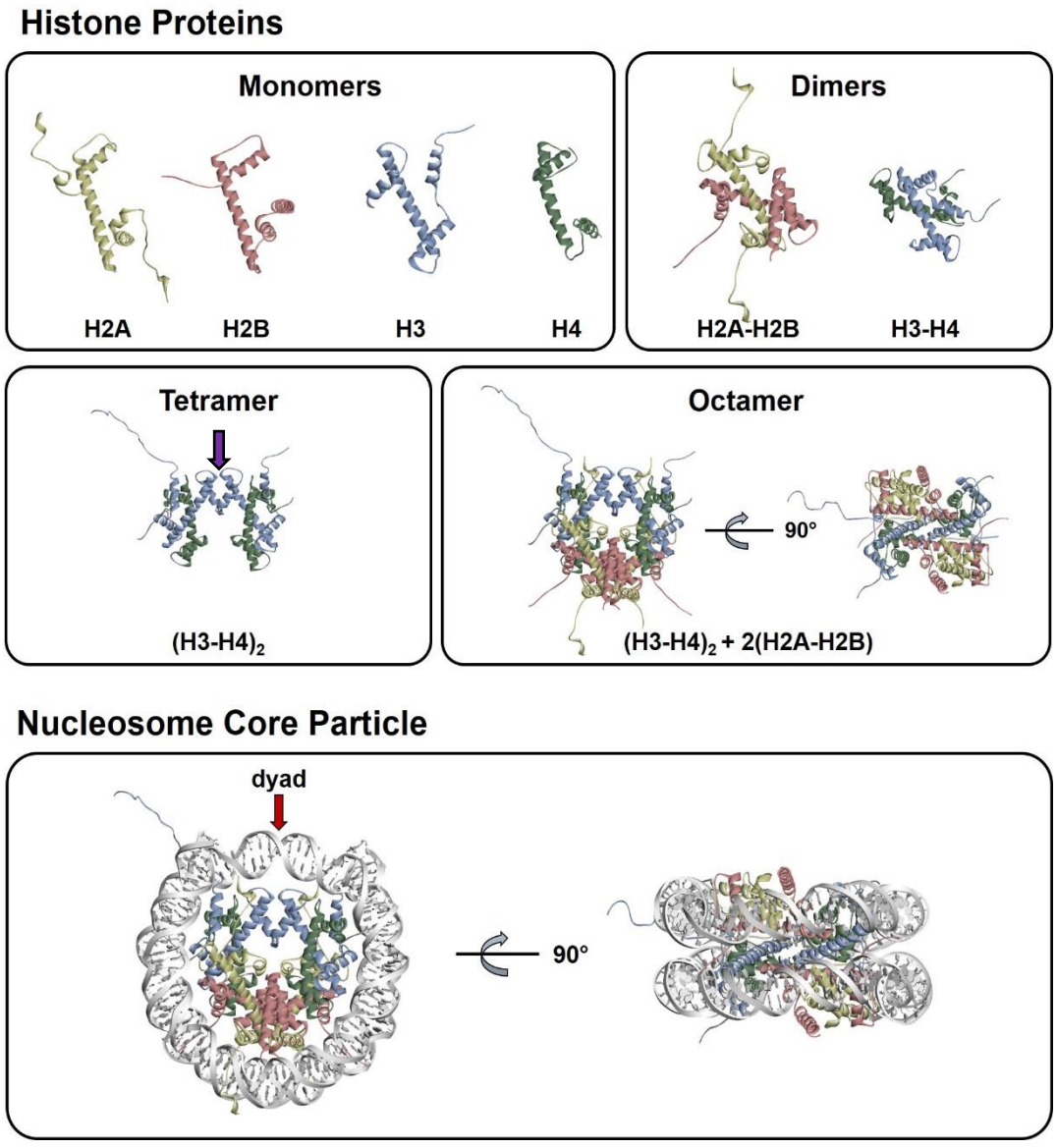


Figure 1.2 The hierarchical organization of nucleosomes
(*Upper*) Hierarchical assembly of core histone proteins. The four monomer species share the same histone fold: a long alpha helix flanked by two short alpha helices. All four histones have flexible N-terminal tails (not completely resolved), while H2A has an additional C-terminal tail. Heterodimer pairs of H2A-H2B and H3-H4 are formed though the "handshake" motif as shown. Two H3-H4 heterodimers dimerize through a very stable four-helix bundle (purple arrow) to form a tetramer. An H2A-H2B dimer interacts with both ends of the tetramer to form the full octameric histone core.
(*Lower*) Crystal structure of the nucleosome core particle with 146 bp DNA.
PDB model used is from [50]

Significant portions of the histone proteins (nearly 30% of the mass) are unresolved in the crystal structures due to their inherent flexibility [50,61]. These regions are thought to be unstructured and are called "histone tails" since they extend from the N-terminus of the histone proteins (though H2A also has a C-terminal histone tail). Basic residues on the histone tails align them along the minor grooves of the DNA and guide them towards the exterior surface of the NCP. This exposure makes histone tails accessible to kinases [62], methyltransferase [63], acetyltransferase [64], and other chemical modifications [65,66] that are used to modulate nucleosome structure and function. Sometimes referred to as "master control switches," histone tails not only regulate DNA accessibility within individual nucleosomes [67], but also facilitate inter-nucleosome interactions and direct how nucleosome arrays are folded [48]. Evidence for this latter role is observed within the crystal structure itself, where an H4 tail from the adjacent nucleosome in the crystal makes direct contact with an acidic patch on the surface of the nucleosome [50].

The stability of the NCP structure is maintained through electrostatic contacts between the DNA and histone proteins. With a negative charge provided by each phosphate group in its backbone, DNA is a highly charged and self-repulsive polymer that is relatively rigid on the length scale of the nucleosome. At near physiological ionic strength (≈ 150 mM NaCl), DNA has a persistence length (≈ 500 Å, ≈ 150 base-pairs) that is significantly longer than the superhelical diameter of the nucleosome (≈ 100 Å) [68]. Since the persistence length represents the distance across which a polymer resists bending due to thermal fluctuations alone, significant energy must be spent (relative to k_BT) to bend the DNA into a nucleosome [69].

Basic residues on the surface of the histone core interact with the negatively charged backbone of the DNA to neutralize the charge and bend the DNA. The non-specificity of this interaction (via phosphates in the backbone, as opposed to the bases) allow histones to organize virtually any DNA sequence. However, the mechanical properties of the DNA play important roles in determining the energy cost to bend DNA [69]. The inclusion of specific motifs, such as the periodic placement of TA dinucleotide steps every 10 base-pairs, has been shown to lower DNA bending energy and favor nucleosome formation for sequences containing these motifs [70,71]. One notable example of a sequence with mechanical properties that are conducive for nucleosome formation is the "601 DNA." This sequence was engineered using SELEX (systematic evolution of ligands by exponential enrichment) to bind very tightly to histones, and is commonly used in biophysical research (including this work) to produce homogeneously positioned nucleosomes [71].

Although the nucleosome core particle is an indispensable source of insight into the interactions that stabilize nucleosomes, nucleosomes form a range of distinct structures that vary in both conformation and composition. To date, all crystallographic structures of mono-nucleosomes contain between 145-147 base-pairs of DNA [50,61,72]. The specificity of this range reflects the importance of crystal contacts formed by base-stacking interactions between the DNA ends of neighboring nucleosomes. Hence in solution, the trajectories of the last 10-20 base-pairs may vary. Early nucleosome studies with octameric histone cores have been observed to stably accommodate anywhere from ≈ 100-170 base-pairs of DNA [73–76], suggesting that nucleosomes may exhibit alternate conformations in addition to that observed in the NCP.

1.3 Alternate nucleosomes and dynamics

Recent studies demonstrate strong evidence for nucleosome dynamics. Experiments testing the DNA accessibility of restriction endonucleases on NCPs reveal that sites along the entire DNA are transiently exposed [77,78]. This exposure is more prevalent near the DNA ends and decreases towards the dyad axis. Single-particle FRET experiments reveal that the DNA ends spontaneously unwrap and rewrap on timescales of 100-250 ms [35,79,80]. Such motions, termed "breathing," are proposed to be an important mechanism for DNA-binding proteins to gain initial access to sequences sequestered in the NCP. The higher propensity for the DNA ends to unwrap is consistent with optical tweezer experiments that show multi-stage DNA release [81,82]. Using force induced disassembly, lower forces were required to unwind the outer DNA (associated with the H2A-H2B dimers), compared to the DNA closer to the dyad (associated with the $(H3\text{-}H4)_2$ tetramer).

Nucleosomes with eight histone proteins, referred to as octasomes, may not be the only biologically relevant composition of this histone core. Nucleosomes missing one (hexasomes) [83–86] or two H2A-H2B dimers (tetrasomes) [87,88] have been reported. The biological significance of these alternate forms is intriguing, as they may serve as key intermediate conformations during gene expression and nucleosome assembly/disassembly. Since hexasomes break the two-fold symmetry of canonical nucleosomes, they may be important intermediates in directing asymmetric activity on nucleosomes. For example, the hexasome has been suggested as the active form of the NCP during transcription [85]. Two general mechanisms have been proposed for hexasome formation: (1) by direct invasion by proteins (as implicated in the activity of

polymerase [85], histone chaperones [45] and chromatin remodelers [89]), or (2) through the application of superhelical torsion on topologically constrained DNA. The creation of positive superhelical torsion, which is naturally induced in front of moving polymerase, is sufficient to evict one H2A-H2B dimer from the NCP [90].

The modular nature of the core histones facilitates the formation and stabilization of alternate nucleosome configurations. By incorporating different combinations of histone variants and post-translational modifications, nucleosomes have numerous opportunities to diversity in structure and function. The formation and displacement of these variants are driven through the activity of histone chaperones [91,92] and chromatin remodelers [57,60]. Furthermore, the modular architecture may enable alternate histone core conformations such as the "open intermediate" [93] and "split nucleosome" [94,95]. In the former intermediate, the interfaces between H2A-H2B dimer and $(H3\text{-}H4)_2$ tetramer are disrupted; whereas in the latter, the nucleosome is split into two halves containing (H3-H4)-(H2A-H2B). Though the prevalence and biological relevance of these forms are yet to be determined, these conformations expose new protein surfaces and provide new opportunities for interactions with nuclear proteins.

1.4 Summary of research

The central aim of this work is to gain structural insights into alternate nucleosome configurations that may serve as key intermediates in nucleosome processing. Since the large size and inherent conformational flexibility of partially unwrapped nucleosomes present challenges for structural determination using traditional techniques such as x-ray crystallography and NMR, we applied a novel approach based on small angle x-ray

scattering (SAXS) and ensemble modeling to characterize alternate nucleosome conformations. An outline of this dissertation is as follows.

In Chapter 2, we introduce the principal experimental strategies applied in this work: (1) contrast variation with small angle x-ray scattering (CV-SAXS), and (2) ensemble modeling. We first describe the theoretical background behind small angle x-ray scattering and explain how contrast variation methods can be applied to glean insights into nucleic acid conformations within protein-nucleic acid systems. We then demonstrate how the pairing of CV-SAXS with ensemble modeling provides new opportunities to resolve DNA structures and overcome the typical limitations in interpreting SAXS data.

In Chapters 3-5, we present three separate studies that apply these approaches to investigate the formation of alternate nucleosome conformations. These studies demonstrate progressive advancements in the experimental strategy as they address deeper and more biologically relevant questions.

In Chapter 3, we describe the first application of this approach to monitor the structures of DNA during the salt-induced disassembly of nucleosomes [96]. Following a rapid salt jump to a condition that destabilizes nucleosomes (2.0 M NaCl), we observe the transient formation of an asymmetrically unwrapped DNA structure. The persistence of this structure for several hundred milliseconds hints that nucleosomes may be predisposed to unwrap asymmetrically and that this structure may be the preferred substrate for other proteins that interact with the nucleosome. Notably, this was the first SAXS study to demonstrate contrast variation while achieving time-resolution on the 100 ms timescale.

In Chapter 4, we build upon the previous study and investigate the formation pathways of the asymmetric intermediate, and relate the changes in DNA conformation with the composition of the histone core [97]. By jumping to a lower salt concentration (1.2 M NaCl), we limit the speed and extent of disassembly, and resolve two pathways through which the asymmetric intermediate is formed. Using time-resolved FRET to monitor the release of H2A-H2B dimers, we observe sequential dimer release that follows the asymmetric release of the DNA. This link between DNA conformation and histone configuration has intriguing implications on how proteins may remodel nucleosomes.

In Chapter 5, we shift towards a more biological system and explore how the chromatin remodeler Chd1 (chromodomain-helicase DNA-binding protein 1) alters nucleosome structure. In response to Chd1 binding, DNA is partially unwrapped along the outer turn of the nucleosome (≈ 45 base-pairs from the dyad axis). However, the trajectory of the unwrapped DNA varies upon the addition of different ATP-analogs (ADP-BeFx, AMP-PNP). These results suggest that the conformational states of the ATPase domain play an important role driving the activity of Chd1.

REFERENCES

[1] J.D. Watson, F.H.C. Crick, Molecular Structure of Nucleic Acids, Nature. (1953) 737–738.

[2] G.M. Church, Y. Gao, S. Kosuri, Next-Generation Digital Information Storage in DNA, Science (80-.). 1226355 (2012) 1–2.

[3] K.E. van Holde, Chromatin, Springer S, Springer-Verlag, 1989.

[4] J.R. Dixon, I. Jung, S. Selvaraj, Y. Shen, J.E. Antosiewicz-Bourget, A.Y. Lee, Z. Ye, A. Kim, N. Rajagopal, W. Xie, Y. Diao, J. Liang, H. Zhao, V. V. Lobanenkov, J.R. Ecker, J.A. Thomson, B. Ren, Chromatin architecture reorganization during stem cell differentiation, Nature. 518 (2015) 331–336.

[5] G. Felsenfeld, M. Groudine, Controlling the double helix, Nature. 421 (2003) 448–453.

[6] P. Fransz, H. De Jong, From nucleosome to chromosome: A dynamic organization of genetic information, Plant J. 66 (2011) 4–17.

[7] C.L. Woodcock, R.P. Ghosh, Chromatin Higher-order Structure and Dynamics, 2 (2010) 1–26.

[8] Y. Lorch, B. Maier-Davis, R.D. Kornberg, Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions, Genes Dev. 28 (2014) 2492–2497.

[9] S. Balasubramanian, F. Xu, W.K. Olson, DNA sequence-directed organization of chromatin: Structure-based computational analysis of nucleosome-binding sequences, Biophys. J. 96 (2009) 2245–2260.

[10] B. Suter, Poly(dA-dT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters in vivo, Nucleic Acids Res. 28 (2000) 4083–4089.

[11] K. Sarma, D. Reinberg, Histone variants meet their match, Nat. Rev. Mol. Cell

Biol. 6 (2005) 139–149.

[12] P.B. Talbert, S. Henikoff, Histone variants — ancient wrap artists of the epigenome, Nat. Rev. Mol. Cell Biol. 11 (2010) 264–275.

[13] Y.J. Park, P.N. Dyer, D.J. Tremethick, K. Luger, A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome, J. Biol. Chem. 279 (2004) 24274–24282.

[14] B. Li, M. Carey, J.L. Workman, The Role of Chromatin during Transcription, Cell. 128 (2007) 707–719.

[15] C.L. Peterson, M.-A. Laniel, Histones and histone modifications, Curr. Biol. 14 (2004) R546–R551.

[16] J.S. Choy, S. Wei, J.Y. Lee, S. Tan, S. Chu, T.H. Lee, DNA methylation increases nucleosome compaction and rigidity, J. Am. Chem. Soc. 132 (2010) 1782–1783.

[17] J.Y. Lee, T.H. Lee, Effects of DNA methylation on the structure of nucleosomes, J. Am. Chem. Soc. 134 (2012) 173–175.

[18] S.A. Grigoryev, Nucleosome spacing and chromatin higher-order folding, Nucleus. 3 (2012) 493–499.

[19] J. Bednar, R.A. Horowitz, S.A. Grigoryev, L.M. Carruthers, J.C. Hansen, A.J. Koster, C.L. Woodcock, Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin, Proc. Natl. Acad. Sci. 95 (1998) 14173–14178.

[20] P. Zhu, G. Li, Structural insights of nucleosome and the 30-nm chromatin fiber, Curr. Opin. Struct. Biol. 36 (2016) 106–115.

[21] S. A. Grigoryev, C.L. Woodcock, Chromatin organization - The 30nm fiber, Exp. Cell Res. 318 (2012) 1448–1455.

[22] K. Maeshima, S. Hihara, M. Eltsov, Chromatin structure: Does the 30-nm fibre exist in vivo?, Curr. Opin. Cell Biol. 22 (2010) 291–297.

[23] J.T. Finch, A. Klug, Solenoidal model for superstructure in chromatin., Proc. Natl. Acad. Sci. U. S. A. 73 (1976) 1897–901.

[24] B. Dorigo, Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber, Science (80-.). 306 (2004) 1571–1573.

[25] T. Schalch, S. Duda, D.F. Sargent, T.J. Richmond, X-ray structure of a tetranucleosome and its implications for the chromatin fibre, Nature. 436 (2005) 138–141.

[26] C. Wu, A. Bassett, A. Travers, A variable topology for the 30-nm chromatin fibre, EMBO Rep. 8 (2007) 1129–1134.

[27] N.L. Adkins, M. Watts, P.T. Georgel, To the 30-nm chromatin fiber and beyond, Biochim. Biophys. Acta - Gene Struct. Expr. 1677 (2004) 12–23.

[28] J. Mirkovitch, M.E. Mirault, U.K. Laemmli, Organization of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold, Cell. 39 (1984) 223–232.

[29] U.K. Laemmli, E. Kas, L. Poljak, Y. Adachi, Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains, Curr. Opin. Genet. Dev. 2 (1992) 275–285.

[30] K.A. Holland, Chromosomes: Nonhistone Proteins, Encycl. Life Sci. (2002) 1–9.

[31] S. V Razin, Matrix-associated Regions (MARs) and Scaffold Attachment Regions (SARs), Encycl. Life Sci. (2005) 1–3.

[32] J. Chizhong, B.F. Pugh, Nucleosome positioning and gene regulation: advances through genomics, 10 (2009) 161–172.

[33] P. Prinsen, H. Schiessel, Nucleosome stability and accessibility of its DNA to proteins, Biochimie. 92 (2010) 1722–1728.

[34] O. Bell, V.K. Tiwari, N.H. Thomä, D. Schübeler, Determinants and dynamics of genome accessibility., Nat. Rev. Genet. 12 (2011) 554–564.

[35] G. Li, M. Levitus, C. Bustamante, J. Widom, Rapid spontaneous accessibility of nucleosomal DNA., Nat. Struct. Mol. Biol. 12 (2005) 46–53.

[36] A.J. Andrews, K. Luger, Nucleosome structure(s) and stability: variations on a theme., Annu. Rev. Biophys. 40 (2011) 99–117.

[37] D.E. Schones, K. Cui, S. Cuddapah, T. Roh, A. Barski, Z. Wang, G. Wei, K. Zhao, Resource Dynamic Regulation of Nucleosome Positioning in the Human Genome, (2008) 887–898.

[38] C. Hebert, H.R. Crollius, Nucleosome rotational setting is associated with transcriptional regulation in promoters of tissue-specific human genes, (2010).

[39] M. Brand, J.G. Moggs, M. Oulad-abdelghani, J. Stevenin, UV-damaged DNA-binding protein in the TFTC complex links DNA damage recognition to nucleosome acetylation, 20 (2001) 3187–3196.

[40] H. Wang, L. Zhai, J. Xu, H. Joo, S. Jackson, H. Erdjument-bromage, P. Tempst, C. Hill, N. Carolina, Histone H3 and H4 Ubiquitylation by the CUL4-DDB-ROC1 Ubiquitin Ligase Facilitates Cellular Response to DNA Damage, (2006) 383–394.

[41] C.-K. Lee, Y. Shibata, B. Rao, B.D. Strahl, J.D. Lieb, Evidence for nucleosome depletion at active regulatory regions genome-wide, Nat. Genet. 36 (2004) 900–905.

[42] E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thåström, Y. Field, I.K. Moore, J.-P.Z. Wang, J. Widom, A genomic code for nucleosome positioning, Nature. 442 (2006) 772–778.

[43] A. Soufi, M.F. Garcia, A. Jaroszewicz, N. Osman, M. Pellegrini, K.S. Zaret, Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming, Cell. 161 (2015) 555–568.

[44] T. Sekiya, U.M. Muthurajan, K. Luger, A. V. Tulin, K.S. Zaret, Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA, Genes Dev. 23 (2009) 804–809.

[45] R. Belotserkovskaya, FACT Facilitates Transcription-Dependent Nucleosome Alteration, Science (80-.). 301 (2003) 1090–1093.

[46] M.Y. Kim, S. Mauro, N. Gévry, J.T. Lis, W.L. Kraus, NAD+-dependent modulation of chromatin structure and transcription by nucleosome binding properties of PARP-1, Cell. 119 (2004) 803–814.

[47] M. Grunstein, Histone acetylation in chromatin structure and transcription., Nature. 389 (1997) 349–352.

[48] C. Zheng, C. Zheng, J.J. Hayes, J.J. Hayes, Structures and interactions of the core histone tail domains., Biopolymers. 68 (2003) 539–46.

[49] P.-Y. Kan, T.L. Caterino, J.J. Hayes, The H4 Tail Domain Participates in Intra- and Internucleosome Interactions with Protein and DNA during Folding and Oligomerization of Nucleosome Arrays, Mol. Cell. Biol. 29 (2009) 538–546.

[50] K. Luger, A.W. Mäder, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution., Nature. 389 (1997) 251–260.

[51] C.A. Davey, D.F. Sargent, K. Luger, A.W. Maeder, T.J. Richmond, Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution, J. Mol. Biol. 319 (2002) 1097–1113.

[52] R.D. Kornberg, Y. Lorch, Twenty-five years of the nucleosome, fundamental particle of the eukaryotic chromosome., Cell. 98 (1999) 285–294.

[53] S.E. Polo, G. Almouzni, Chromatin assembly: a basic recipe with various flavours., Curr. Opin. Genet. Dev. 16 (2006) 104–11.

[54] R.J. Burgess, Z. Zhang, Histone chaperones in nucleosome assembly and human disease, Nat. Struct. Mol. Biol. 20 (2013) 14–22.

[55] A.J. Andrews, X. Chen, A. Zevin, L. a. Stargell, K. Luger, The Histone Chaperone Nap1 Promotes Nucleosome Assembly by Eliminating Nonnucleosomal Histone DNA Interactions, Mol. Cell. 37 (2010) 834–842.

[56] S. D'Arcy, K. Martin, T. Panchenko, X. Chen, S. Bergeron, L. Stargell, B. Black, K. Luger, Chaperone Nap1 Shields Histone Surfaces Used in a Nucleosome and Can Put H2A-H2B in an Unconventional Tetrameric Form, Mol. Cell. 51 (2013) 662–677.

[57] C.R. Clapier, B.R. Cairns, The Biology of Chromatin Remodeling Complexes, Annu. Rev. Biochem. 78 (2009) 273–304.

[58] R. Sundaramoorthy, A.L. Hughes, V. Singh, N. Wiechens, D.P. Ryan, H. El-Mkami, M. Petoukhov, D.I. Svergun, B. Treutlein, S. Quack, M. Fischer, J. Michaelis, B. Böttcher, D.G. Norman, T. Owen-Hughes, Structural reorganization of the chromatin remodeling enzyme Chd1 upon engagement with nucleosomes, Elife. 6 (2017).

[59] G.D. Bowman, Mechanisms of ATP-dependent nucleosome sliding, Curr. Opin. Struct. Biol. 20 (2010) 73–81.

[60] G. Hauk, J.N. McKnight, I.M. Nodelman, G.D. Bowman, The Chromodomains of the Chd1 Chromatin Remodeler Regulate DNA Access to the ATPase Motor, Mol. Cell. 39 (2010) 711–723.

[61] T.J. Richmond, C.A. Davey, The structure of DNA in the nucleosome core., Nature. 423 (2003) 145–50.

[62] B.D. Strahl, C.D. Allis, The language of covalent histone modifications, 403

(2000) 41–45.

[63] Y. Zhang, D. Reinberg, Transcription regulation by histone methylation : interplay between different covalent modifications of the core histone tails, (2001) 2343–2360.

[64] Y. Sharon, M. John, C. David, Histone acetyltransferases, (2001).

[65] J. Ausio, F. Dong, K.E. Van Holde, Use of Selectively Trypsinized Nucleosome Core Particles to Analyze the Role of the Histone " Tails " in the Stabilization of the Nucleosome, (1989) 451–463.

[66] S.D. Taverna, H. Li, A.J. Ruthenburg, C.D. Allis, D.J. Patel, How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers., Nat. Struct. Mol. Biol. 14 (2007) 1025–1040.

[67] M. Biswas, K. Voltz, J.C. Smith, Role of Histone Tails in Structural Stability of the Nucleosome, 7 (2011) 2–13.

[68] P. Hagerman, Flexibility Of DNA, Annu. Rev. Biophys. Biomol. Struct. 17 (1988) 265–286.

[69] J. Widom, Role of DNA sequence in nucleosome stability and dynamics, Cornell University Library, 2001.

[70] A. Thastrom, P.T. Lowary, H.R. Widlund, H. Cao, M. Kubista, J. Widom, Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences., J. Mol. Biol. 288 (1999) 213–229.

[71] P.T. Lowary, J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning., J. Mol. Biol. 276 (1998) 19–42.

[72] K. Luger, Dynamic nucleosomes., Chromosom. Res. 14 (2006) 5–16.

[73] K. Van Holde, J. Zlatanova, The nucleosome core particle: does it have

structural and physiologic relevance?, BioEssays. 21 (1999) 776–780.

[74] V. V Bakayev, Nucleosomes and Subnucleosomes: Heterogeneity and Composition, Cell. 11 (1977) 619–629.

[75] K. Tatchell, K.E. van Holde, Nucleosome Reconstitution: Effect of DNA Length on Nucleosome Structure, Biochemistry. 18 (1979) 2871–2880.

[76] W.O. Weischet, J.R. Allen, G. Riedel, K.E. van Holde, The effects of salt concentration and H-1 depletion on the digestion of calf thymus chromatin by micrococcal nuclease., Nuclei. 6 (1979) 1843–1862.

[77] K.J. Polach, J. Widom, Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation., J. Mol. Biol. 254 (1995) 130–149.

[78] H.S. Tims, K. Gurunathan, M. Levitus, J. Widom, Dynamics of nucleosome invasion by DNA binding proteins, J. Mol. Biol. 411 (2011) 430–448.

[79] W.J.A. Koopmans, A. Brehm, C. Logie, T. Schmidt, J. Van Noort, Single-pair FRET microscopy reveals mononucleosome dynamics, J. Fluoresc. 17 (2007) 785–795.

[80] M. Tomschik, K. Van Holde, Nucleosome Dynamics as Studied by Single-pair Fluorescence Resonance Energy Transfer: A Reevaluation, J. Fluoresc. 19 (2009) 53–62.

[81] B.D. Brower-Toland, C.L. Smith, R.C. Yeh, J.T. Lis, C.L. Peterson, M.D. Wang, Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA., Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 1960–1965.

[82] M.A. Hall, A. Shundrovsky, L. Bai, R.M. Fulbright, J.T. Lis, M.D. Wang, High-resolution dynamic mapping of histone-DNA interactions in a nucleosome., Nat. Struct. Mol. Biol. 16 (2009) 124–129.

[83] Y. Arimura, H. Tachiwana, T. Oda, M. Sato, H. Kurumizaka, Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome, Biochemistry. 51 (2012) 3302–3309.

[84] J. Zlatanova, T.C. Bishop, J.M. Victor, V. Jackson, K. van Holde, The Nucleosome Family: Dynamic and Growing, Structure. 17 (2009) 160–171.

[85] M.L. Kireeva, W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, V.M. Studitsky, Nucleosome Remodeling Induced by RNA Polymerase II, Mol. Cell. 9 (2002) 541–552.

[86] N. Azegami, K. Saikusa, Y. Todokoro, A. Nagadoi, H. Kurumizaka, Y. Nishimura, S. Akashi, Conclusive Evidence of the Reconstituted Hexasome Proven by Native Mass Spectrometry, 14 (2013) 5155–5157.

[87] A.J. Katan, R. Vlijm, A. Lusser, C. Dekker, Dynamics of Nucleosomal Structures Measured by High-Speed Atomic Force Microscopy., Small. 11 (2014) 1–9.

[88] R. Vlijm, M. Lee, J. Lipfert, A. Lusser, C. Dekker, N.H. Dekker, Nucleosome Assembly Dynamics Involve Spontaneous Fluctuations in the Handedness of Tetrasomes, Cell Rep. 10 (2015) 216–225.

[89] R.F. Levendosky, A. Sabantsev, S. Deindl, G.D. Bowman, The Chd1 chromatin remodeler shifts hexasomes unidirectionally, (2016).

[90] V. Levchenko, B. Jackson, V. Jackson, Histone release during transcription: Displacement of the two H2A-H2B dimers in the nucleosome is dependent on different levels of transcription-induced positive stress, Biochemistry. 44 (2005) 5357–5372.

[91] L.L. Wallrath, M.W. Vitalini, S.C.R. Elgin, Histone Chaperones in the Assembly and Disassembly of Chromatin, 2014.

[92] Z.A. Gurard-Levin, J.-P. Quivy, G. Almouzni, Histone Chaperones: Assisting

Histone Traffic and Nucleosome Dynamics, Annu. Rev. Biochem. 83 (2014) 487–517.

[93] V. Böhm, A.R. Hieb, A.J. Andrews, A. Gansen, A. Rocker, K. Tóth, K. Luger, J. Langowski, Nucleosome accessibility governed by the dimer/tetramer interface, Nucleic Acids Res. 39 (2011) 3093–3102.

[94] M. Lee, W.T. Garrard, Transcription-induced nucleosome "splitting": an underlying structure for DNase I sensitive chromatin, EMB. 10 (1991) 607–615.

[95] M. Lee, W.T. Garrard, Positive DNA supercoiling generates a chromatin conformation characteristic of highly active genes, Proc Natl Acad Sci U S A. 88 (1991) 9675–9679.

[96] Y. Chen, J.M. Tokuda, T. Topping, J.L. Sutton, S.P. Meisburger, S. A. Pabit, L.M. Gloss, L. Pollack, Revealing transient structures of nucleosomes as DNA unwinds, Nucleic Acids Res. 42 (2014) 8767–8776.

[97] Y. Chen, J.M. Tokuda, T. Topping, S.P. Meisburger, S.A. Pabit, L.M. Gloss, L. Pollack, Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core, Proc. Natl. Acad. Sci. 114 (2017) 334–339.

Chapter 2

Methods

Small angle x-ray scattering (SAXS) is a powerful scattering technique that reports the global structure properties of biomolecules in solution [1–4]. The primary strength of SAXS is its ability to probe the unconstrained behavior of biomolecules over a wide range of experimental conditions. This flexibility makes SAXS suitable for investigating dynamics and large conformational changes that are often inaccessible to other techniques. Recent advancements in instrumentation and analysis methods have greatly expanded the role of SAXS in resolving the motions of large biomolecular assemblies [5,6]. High intensity x-ray sources and solution mixing technologies have enabled time-resolved studies to probe dynamics on timescales as short as milliseconds [7,8]. Advancements in ensemble methods now enable interpretation of results beyond the conformational average so conformational heterogeneity can be resolved [9–13].

While the studies in Chapters 3-5 contain their own details on the methods used, here we provide a brief overview of SAXS and the key concepts behind the contrast variation method that enable investigations of DNA conformations within nucleosomes. We then discuss our procedures for ensemble modeling and important considerations for time-resolved SAXS measurements Finally, we present details on the beamline setups used.

Parts of this chapter is adapted with permission from [1]**:** J.M. Tokuda, S.A. Pabit, L. Pollack, Protein-DNA and ion-DNA interactions revealed through contrast variation SAXS, Biophys. Rev. 8 (2016) 139–149.

2.1 Introduction

The basic concept of small angle x-ray scattering is illustrated in Figure 2.1. An aqueous sample is exposed to a collimated and monochromatic x-ray beam (typically 10^{10}-10^{14} photons/s) with a wavelength, λ, between 1-2 Å (≈ 6000-12000 eV). A small fraction of the photons (≈ $1/10^6$) are elastically scattered by the electrons within the illuminated volume (via Thomson scattering) and produce an interference pattern that is imaged onto a detector. Since the molecules that contribute to the signal are sampled in all orientations, SAXS measures the spherically averaged scattering patterns of the molecules in solution. The resulting scattering intensities are radially symmetric in the image, and scattering profiles are generated by first integrating the intensity at each azimuthal angle, and then plotting the intensity, $I(q)$, as a function of the angle, $q = 4\pi \sin(\theta)\,/\lambda$. This definition of the scattering angle (q) makes it independent of the x-ray energy. For most biomolecules, with sizes between 10-100 nm, the measured intensities decay within a few degrees ($2\theta < 5°$). In order to isolate the scattering from the biomolecules alone, the SAXS profile of the buffer (without the biomolecule) is subtracted from the SAXS profile of the sample.

The general equation for the scattering intensity of a dilute, monodisperse particle is given by:

$$I(q) = 4\pi \int_0^{D_{max}} r^2\, \gamma(r) \frac{\sin(q \cdot r)}{q \cdot r} dr\,, \qquad (Eq.\ 2.1)$$

where D_{max} is the maximum dimension of the particle and $\gamma(r)$ is the spherically averaged autocorrelation function of the particle's electron density. Solutions to Eq. 2.1

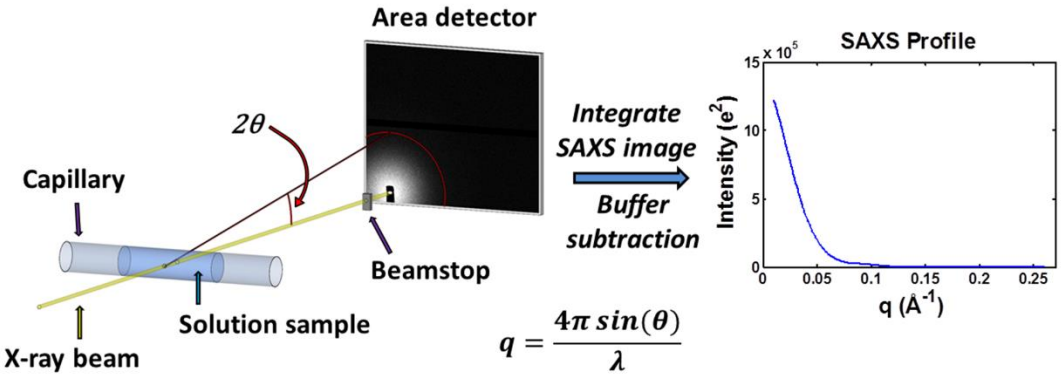


Figure 2.1 SAXS schematic. (*Left*) A schematic of a typical SAXS experiment is shown. The sample is typically a buffered solution containing 2 mg/mL of protein, DNA, or protein-DNA complex. This sample oscillates through a quartz capillary to reduce radiation damage from the x-ray beam. The scattered x-rays are imaged onto an area detector while the primary beam is either blocked or greatly attenuated (as shown) by a beamstop. (*Right*) The images are pooled, averaged, and converted into profiles of intensity as a function of scattering vector, *I(q)*, through azimuthal integration. For each sample, a corresponding measurement of the buffer alone is made and the resulting buffer profile is subtracted from the sample profile to obtain the macromolecular SAXS profile. SAXS intensities can be calibrated onto an absolute scale (in units of e^2) through the measurement of water as a standard. Figure adapted from [1]

are complex and depend on the shape and internal structure of the particle. However, the low angle limit provides some useful information. At the forward scattering angle, $q = 0$, maximal constructive interference is achieved since all of the scattered x-rays are in phase. Thus, the scattering, $I(0)$, depends only on the number of electrons and is related to the molecular mass and number of biomolecules that give rise to the scattering. Using $I(0)$, molecular mass can be determined through comparison with a reference molecule with known mass, or through calibration of the scattering intensities onto an absolute scale using a known standard (e.g. water scattering) [14]. Since the measured scattering from N non-interacting particles is simply the sum of their

scattering, molecular mass measurements require that the concentration of the biomolecule is accurately known. Because the incident x-ray beam is blocked by a beamstop, $I(0)$ is not measured directly and must be extrapolated.

At low angles, the scattering can be approximated as a Gaussian through the Guinier formula [4,15]:

$$I(q) = I(0) \cdot e^{-\frac{(R_g \cdot q)^2}{3}}, \quad (Eq.\ 2.2)$$

where R_g is the radius of gyration. Mathematically, R_g is the root mean square distance between all electron pairs within a particle and is a quantitative measure of size. A linear fit to experimental data plotted as $ln\big(I(q)\big)$ versus q^2 provides $I(0)$ and R_g from the intercept and slope, respectively. For globular structures, this approximation is valid for $q \cdot R_g < 1.3$.

2.2 Scattering theory and contrast variation

The theoretical principles behind SAXS are covered in great detail in other reviews [4,16,17]. Here, we provide only the relevant introductory information to understand the contrast variation methods applied in this work.

We begin by briefly discussing the origin of the SAXS signal (for more details see [4,18]). The net scattering intensity (measured above the background) originates from electron density fluctuations within the illuminated volume. For proteins and nucleic acids in an aqueous buffer, these fluctuations are dominated by the higher electron densities of the protein (≈ 420 e^-/nm^3) and nucleic acids (≈ 550 e^-/nm^3) over the surrounding solution (≈ 334 e^-/nm^3) [2]. The average electron density difference (or "contrast") between a macromolecule and its surrounding solution can be expressed as:

$$\Delta\rho_M = \rho_M - \rho_{sol} \,, \quad (Eq.\ 2.3)$$

where ρ_M and ρ_{sol} are the average electron density values for the macromolecule and solution, respectively. The number of excess electrons present in the volume occupied by the macromolecules (relative to an identical volume occupied by the solvent) is given by the scattering factor:

$$f_M = \Delta\rho_M \cdot V_M \,, \quad (Eq.\ 2.4)$$

where V_M is the volume of the macromolecules. If the macromolecules were in vacuum, f_M would equal the number of electrons present in the macromolecules. Since the solution is essential for physiological measurements, the relevant value reflects the number of excess electrons above that contained by the solution in the same volume.

The amplitude of the scattering signal is described by the product of this scattering factor, f_M, and an angle (or q)-dependent form factor, $F_M(q)$, that reflects the arrangement of electrons in the macromolecule. For a single component system, the scattering amplitude is given by:

$$A = f_M \cdot F_M(q) \,, \quad (Eq.\ 2.5)$$

This equation neglects the internal density fluctuations of the macromolecule [2]. SAXS experiments measure intensity, the product of the scattering amplitude and its complex conjugate: $A \cdot A^*$. For a single particle system, this "squaring" is straightforward. The measured scattering intensity can be written as:

$$I(q) = {f_M}^2 \cdot P_M(q) \,. \quad (Eq.\ 2.6)$$

Here, $P_M(q)$ is the partial scattering form factor of this macromolecule, given by $F_M(q){F_M}^*(q)$ integrated over all space. $P_M(q)$ reflects the shape of the macromolecule

and (by convention) is unity at $q = 0$, hence $I(0)$ is related to the number of excess electrons squared.

For a two-component system consisting of DNA and protein, the total scattering amplitude is a sum over all components:

$$A = f_{DNA}F_{DNA}(q) + f_{prot}F_{prot}(q)\ , \qquad (Eq.\ 2.7)$$

The resulting expression for scattering intensity contains cross-terms, proportional to various products of the form factors, $F_{DNA}(q)$ and $F_{prot}(q)$. These terms, which depend on the structures of both particles, introduce new challenges for the study of multicomponent systems. For this system, the scattering intensity $I(q)$ is given by:

$$I(q) = {f_{DNA}}^2 P_{DNA}(q) + 2f_{DNA}f_{prot}P_{DNA\cdot prot}(q) + {f_{prot}}^2 P_{prot}(q)\ , \quad (Eq.\ 2.8)$$

where $P_{DNA}(q)$ and $P_{prot}(q)$ are the partial scattering form factors for each of the two components and reflects their individual contributions. The cross-term, $P_{DNA\cdot prot}(q)$, contains information about the relative distributions of electrons between the two components.

As seen in *Eq. 2.8*, SAXS profiles for DNA-protein complexes consist of contributions from the shapes and relative distributions of the individual components. Deconvolution of the signal into the contributions from each component is non-trivial, and this ambiguity severely limits how SAXS profiles can be interpreted.

Contrast variation describes the manipulation of the scattering factors ($f_{DNA}\ and\ f_{prot}$) in the above equations. As seen from *Eq. 1* and *Eq. 2*, the scattering factors are directly related to electron densities: $f_{DNA} \propto \rho_{DNA} - \rho_{sol}$ and $f_{prot} \propto \rho_{prot} - \rho_{sol}$. If the solution electron density (ρ_{sol}) is increased to equal that of the

protein (ρ_{prot}), then $f_{prot} = 0$, and the second and third terms of *Eq. 2.8* vanish (note: *Eq. 2.8* becomes *Eq. 2.6*, which describes a single component system). This is referred to as the "contrast matched" condition and must be found experimentally. The addition of 50%-65% (w/v) sucrose to the solvent is sufficient to match the electron density of most proteins. As stated above, this analysis does not account for internal density fluctuations within the DNA or protein components — a valid approximation in most cases since these fluctuations are typically much smaller than the differences between the components. In the contrast matched condition, the resulting scattering profile is dominated by the DNA shape, and conformational details of the DNA emerge that are otherwise obscured in standard SAXS measurements. Note that the effective scattering factor for the DNA is also reduced when the solution electron density is increased, which results in a weaker but otherwise unchanged signal. This approach is illustrated in Figure 2.2.

Theoretically, if the electron density of the solvent was increased further to match that of the DNA, the protein signal from the complex can be isolated. Sucrose cannot be used to achieve this matching condition due to solubility limits; however, the use of other highly soluble additives with heavier atoms may enable such studies in the future. Small angle neutron scattering (SANS) is better suited for this type of contrast matching due to the remarkable difference in scattering between hydrogen and deuterium. In SANS, the scattering from various components can be selected by varying the H2O/D2O composition of the solvent. This approach has been applied with great success in the pioneering studies of nucleosomes [19–21] as well as a variety of other

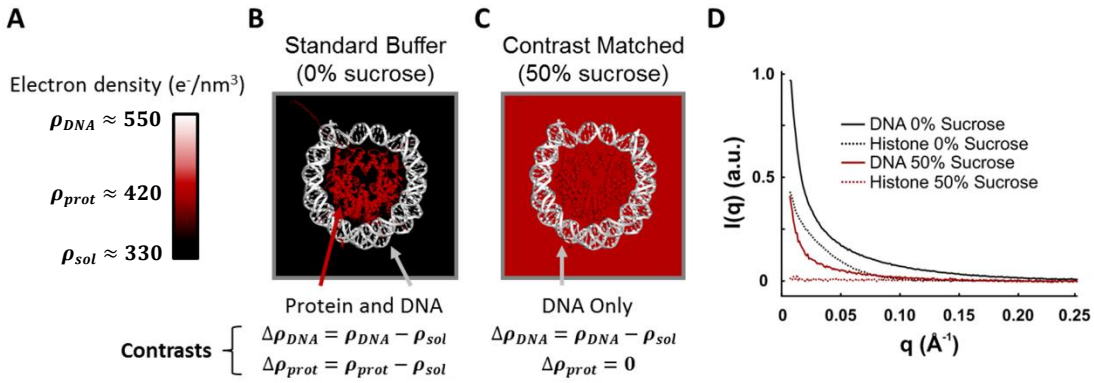


Figure 2.2 Cartoon illustration of contrast variation. (**A**) A color scale bar is shown with average electron density values for DNA, protein, and water. (**B-C**) Cartoon models of the nucleosome core particle (PDB: 1AOI [36]) in buffers with electron densities that vary according to the presence of 0% (**B**) and 50% (**C**) sucrose. The resulting contrasts (excess electron densities) are shown below each condition. (**D**) SAXS profiles for the DNA and histone proteins measured separately with and without sucrose. Note: in 50% sucrose, the histone SAXS signal "disappears," but the DNA is still visible due to its higher electron density.
Figure adapted from [1]

biomolecular complexes [22–27]. Contrast variation SAXS and SANS methods are reviewed in [26–29].

2.3 Applying CV-SAXS to NCPs

The first step in a contrast variation experiment is to determine how much sucrose needs to be added to increase the density of the buffer until it matches that of the proteins (Figure 2.2C). To accomplish this, SAXS profiles of the octameric histone proteins were acquired in solutions containing different concentrations of sucrose. In 50% (w/v) sucrose, the signal from the histone proteins was effectively eliminated (Figure 2.2D). Importantly, SAXS profiles of the DNA acquired in 50% sucrose show that sufficient signal remains for analysis.

The benefits of contrast variation becomes clear when comparing SAXS data for

wrapped NCPs with and without sucrose. For a more intuitive perspective of the data, we can first compute the pairwise distance distribution function, $P(R)$, through a Fourier transform of the SAXS profile as shown below [16]:

$$P(R) = \frac{1}{2\pi^2}\int_0^\infty I(q)\cdot qR\cdot \sin(qR)\,dq\ . \qquad \text{(Eq. 2.9)}$$

This procedure allows conversion of the SAXS data from reciprocal- (q) to real-space (R), and can be very informative if well-defined shapes are present. In practice, only approximate solutions for $P(R)$ are obtained through indirect Fourier transform methods due to experimental limitations (e.g. finite measurable q-range). Detailed instructions for obtaining reliable $P(R)$ curves are reviewed in [6], and $P(R)$s are relatively straightforward to compute using publicly available software such as GNOM [30]. Peaks in the $P(R)$ curves represent length scales that are prominent or repeated in the particles, and assist in interpreting structural changes reported by SAXS. The expected DNA structure and theoretical $P(R)$ for wrapped NCPs are shown in Figure 2.3A-B. Characteristic dimensions of the NCP structure, such as the distance between overlapping DNA ends, and the diameter of the superhelical coil, appear as peaks in $P(R)$ (Figure 2.3A-B). These features are obscured in the $P(R)$ curve for NCPs measured without sucrose since both the protein and DNA contribute to the signal (Figure 2.3C). The bell-shaped $P(R)$ curve observed without sucrose reflects a globular structure and conformational insight is limited to global properties such as the radius and maximum dimension. However, when the NCP is measured with 50% sucrose, the protein scattering is effectively eliminated and identifiable features emerge from the $P(R)$ (Figure 2.3C) that correspond with what is expected for the wrapped NCP structure (Figure 2.3A-B).

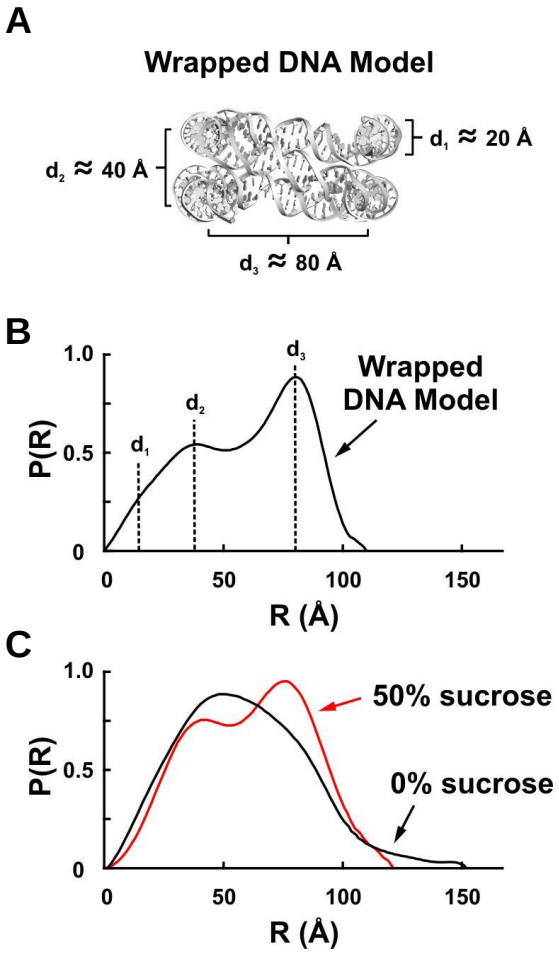


Figure 2.3 Comparison of the structural insight provided by the pair-wise distance distribution ($P(R)$) for NCPs with and without. (**A**) DNA model for the completely wrapped NCP (PDB: 1AOI, [36]). (**B**) Theoretical P(R) curve for the DNA model shown in (**A**). Peaks in the P(R) curve are associated with structural features as follows: d_1 – diameter of the duplex DNA; d_2 – distance between overlapping DNA ends; d_3 – diameter of the overall wrapped structure. (**C**) Comparison of P(R) curves for NCPs measured in 0.2 M NaCl and 0% or 50% sucrose. The differences between the measured P(R)s in (**C**) illustrate how structural details emerge in the contrast matched condition.
Figure adapted from [1]

2.4 Ensemble modeling

Structural modeling provides new opportunities to expand the scope and depth of structural insights provided by SAXS. In traditional SAXS analysis, properties such as molecular mass, radius of gyration, P(R), and low-resolution shape envelopes are

computed directly from SAXS profiles. The development of efficient and reliable programs that calculate theoretical SAXS profiles from PDB models, such as CRYSOL [31], allow direct comparison of experimental data with theoretical scattering profiles from high-resolution structures. Provided a set of plausible atomistic models, SAXS can be combined with refinement algorithms to distinguish between competing models generated by other structural techniques such as x-ray crystallography, NMR, cryo-EM, and molecular dynamics [32]. For example, the unresolved regions in a crystal structure can be modeled using molecular dynamics to generate a pool of possible conformations, and SAXS can be applied as a constraint to select the most plausible model [33,34].

Since SAXS is sensitive to the full ensemble of populated conformations, ensemble strategies provide new opportunities to resolve conformational heterogeneity beyond the conformational average. For non-interacting particles, the experimentally measured SAXS profile is a linear combination of the SAXS profiles for the individual conformations present in solution:

$$I(q) = \sum_{i=1}^{N} {f_i}^2 \cdot P_i(q) \,, \qquad \text{(Eq. 2.10)}$$

where f_i is the scattering factor, and $P_i(q)$ is the partial scattering form factor of the i-th specie. Here, the scattering factor is $f_i = n_i \cdot \Delta\rho_i$, where n_i is the number of times that conformation is represented, and $\Delta\rho_i$ is the number of excess electrons for the conformation. In the case that the conformations share the same number of excess electrons, the measured SAXS profile is simply the weighted sum of the SAXS profiles of the individual conformations. The goal of ensemble strategies is to find a subset of models from a large pool of possible conformations that best recapitulates the SAXS data. Details on how the DNA pools were generated are described in each chapter.

For the work described in the following chapters, we apply an ensemble method adapted from the ensemble optimization method (EOM) developed by Bernado et al for characterizing intrinsically flexible proteins [9–11]. The ensemble method consists of three main steps: (1) generation of a library of conformations, (2) calculation of SAXS profiles for each conformation, (3) application of a search algorithm to find a combination of conformations that best fits the experimental SAXS profile.

Since the library of conformations can be considerably large (≈ 10,000) and each conformation can be represented with different weights, significant computational resources must be applied if every possible combination is to be examined. We apply a genetic algorithm to search through various combinations until the quality of the fits converges [9–11]. The genetic algorithm works as follows: (1) For the first generation, an initial set of K ensembles ("chromosomes") are generated, and each ensemble is populated with 20 random conformations ("genes"). For this explanation, we will assume $K = 50$. (2) For the second generation, $3 \cdot K$ (or 150) new ensembles are generated containing various degrees of cross-over mutations (genes randomly swapped between chromosomes) and random mutations (inclusion of random conformations from the pool). (3) The resulting SAXS profile for each ensemble is computed and ensembles are ranked according to the χ^2-fits to the experimental data. (4) The K (or 50) best ensembles are propagated to the next generation and the evolution process (steps 2-3) is repeated through 10,000 generations. In most cases, convergence is achieved after several thousand generations, resulting in ensembles that fit the data with χ^2-values near 1.0. The entire algorithm is iterated from 10 to 100 times to obtain statistics, with each iteration starting from a random generation of ensembles. This

ensures that the final representative ensemble is reproducibly selected. In the case that similar, but not identical structures, are selected in the final ensembles, the results for all iterations are pooled and similar structures are binned according to RMSD. For NCP studies, the final ensemble typically consists of 2-5 representative structures. The ensemble method is illustrated in Figure 2.4.

2.5 Time-resolved SAXS

The combination of SAXS and mixer technologies is a powerful strategy for investigating conformational dynamics. When the solution environment is rapidly altered, SAXS can be used to measure time-resolved structural changes in response to variations in pH, ionic strength, or the addition of ligands.

The time-resolved SAXS studies in chapters 3-4 use a commercial stopped flow mixer (SFM-400, BioLogic, France) to study the conformational response of NCPs induced by [NaCl] jumps (Figure 2.5). In stopped flow mixing, solutions are directed through a turbulent mixer and into an observation flow cell. The mixed fluids are rapidly stopped by the closure of a solenoid valve and SAXS data are collected thereafter. The earliest time point accessible using a stopped flow mixer (SFM) is limited by the "dead time" of the mixer— the time required for the samples to travel from the mixing region to the observation cell. This delay typically ranges from 1-5 ms depending on the sample flow rate and volume between these two regions. Since SAXS data are recorded in real-time (after mixing), the frame rate of the detector also limits the time-resolution. The Pilatus detector has an image readout time of 3 ms, which means frames can be acquired every 4 ms (1 ms collection + 3 ms readout). As a result,

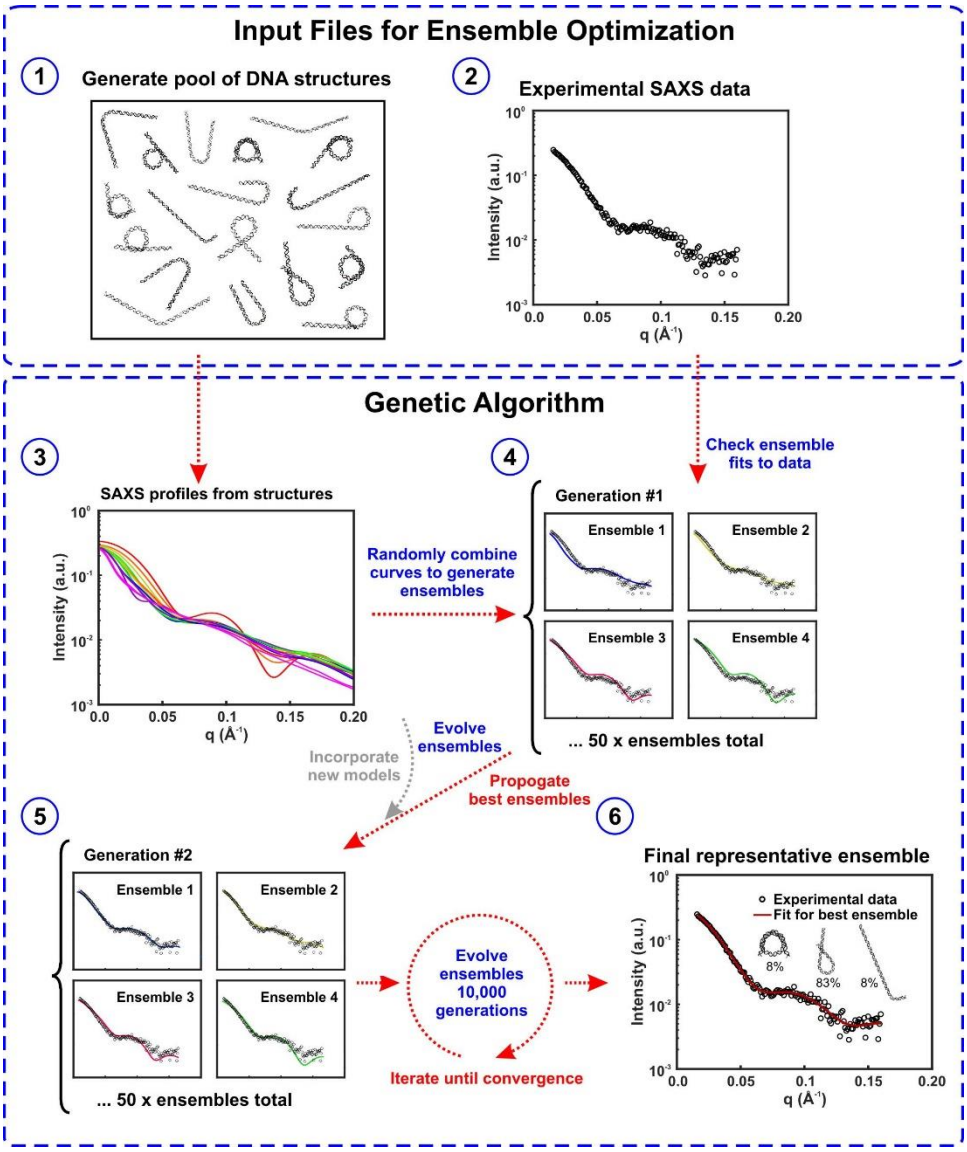


Figure 2.4 Schematic of ensemble optimization method applied to find an ensemble of DNA structures that best fit the data. The procedures are as follows: (1) Generate a large pool of DNA structures to cover the conformational landscape. (2) Collect experimental data for NCPs with sucrose. (3) Calculate SAXS profiles for all structures in the DNA pool. (4) Randomly assign the first generation of ensembles and compare fits to the data. (5) Evolve the ensembles by propagating structures from the best fitting ensembles and incorporating new structures from the pool. Compare fits for the new generation of ensembles. (6) Repeat the evolution cycle until the convergence is achieved. The best fitting ensemble in the final generation is the representative ensemble.

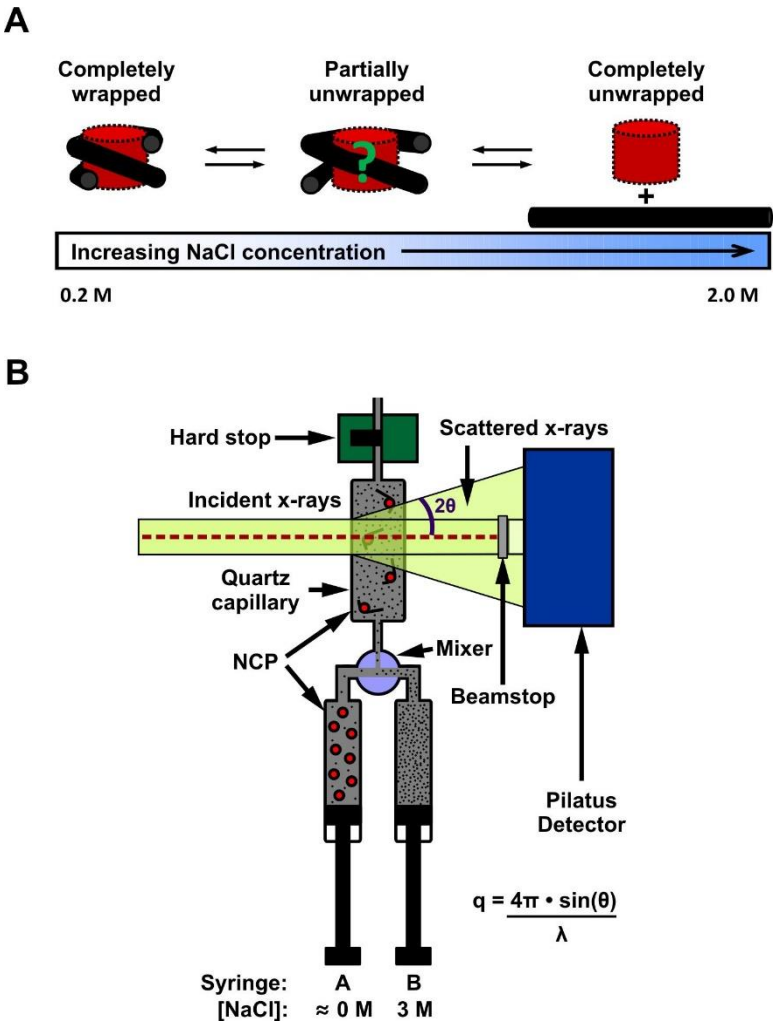


Figure 2.5 Scheme for studying time-resolved salt-induced disassembly of NCPs**.** (**A**) Cartoon showing the use of salt to destabilize NCPs and populate alternate conformations with DNA unwrapping. (**B**) Diagram of stopped flow mixer used for time-resolved SAXS experiments. Intact NCPs in low salt are rapidly mixed with high salt buffers into an observation cell. The flow is stopped by a solenoid valve in the hard stop and the x-ray shutter is immediately opened, exposing the samples to x-rays as the detector records images of the scattered x-rays.

the nominal time resolution achievable using SFM is 9-10 ms. Corresponding images (in time) from multiple measurements are binned for improved statistics. The longest time point accessible using a stopped flow mixer is limited to 30-60 s due to radiation

damage and backflow mixing. The former concern can be addressed through the use of attenuators to reduce x-ray exposure. However, backflow, where the pre-mixed samples diffuse with the mixed samples, is unavoidable and becomes noticeable after 30-60 s.

Efficient mixing of high viscosity solutions containing high sucrose and salt concentrations requires special considerations. Since bubbles are the bane of all mixing experiments, all samples must be degassed and loaded with care. With the SFM mounted vertically, several rounds of aspirating and dispensing should remove most bubbles trapped in the viscous solutions. Nevertheless, the first few shots after sample loading will inevitably be graced by bubbly friends. The commercial SFM used (SFM-400) has four syringes that are independently driven by motors. Since the four syringes are combined sequentially before being directed through the mixer, syringes 1-3 are loaded with the high salt buffer and syringe 4 (the one that is most proximal to the mixer) is loaded with the NCP sample in low salt buffer. This minimizes both dead time and dispersion. In order to achieve the most reproducible mixing and reduce the backflow from the mechanical backlash caused by pressure spikes, all four syringes must be driven for each measurement. A high-density mixer (HDS model) should be used to minimize convection currents that may compromise mixing. One challenge for working with viscous solutions is flushing out the sample from the previous measurement. The key is to use large volumes (> 375 μL) and high flow rates (> 7.5 ml/s). Since the sample cell only holds 30 μL, over 90% of the samples are sacrificed for flushing. Stopped flow mixing performance should be characterized and optimized by fluorescence quenching experiments (e.g. mixing a fluorophore with KI).

2.6 Beamline considerations

Here we describe some of the basic considerations for setting up time-resolved experiments at the G1 station at Cornell High Energy Synchrotron Source (CHESS) (Figure 2.6) and BioCAT Sector 18ID at Advanced Photon Source (APS) (Figure 2.7). For each experiment, the timing must be synchronized between the stopped flow mixer, detector, and x-ray shutter to within 1 ms. Response times for each component must be carefully measured using an oscilloscope and appropriate delay times must be programmed into the software used to coordinate their actions (e.g. SPEC). An external trigger signal from the SFM is used to synchronize the hard-stop closure with the start of data acquisition and x-ray shutter opening.

For experiments conducted at CHESS, the intensity of the incident x-rays decays noticeably over the timescale of a single measurement. In order to compensate for this decay, SAXS curves must be normalized by the intensity of the incident beam. In standard SAXS beamlines, counts of the incident x-rays are usually measured by PIN diodes embedded into the beamstop. However, due to a variety of factors (i.e. electronic noise and timing errors), these counts are unreliable for the short exposures (several ms) required in SFM experiments.

An alternate strategy is to use a semi-transparent beamstop to attenuate and directly image the incident x-ray beam onto the detector. For this purpose, a beamstop can be made from molybdenum (Mo) foil. For x-rays with energies between 10-11 keV, the attenuation length for Mo is $\approx$ 13 μm [35], and Mo beamstops between 100-200 μm (thick) provide sufficient attenuation. The thickness of the foil is chosen so the counts in every pixel remain within their dynamic range during the exposure. For the Pilatus

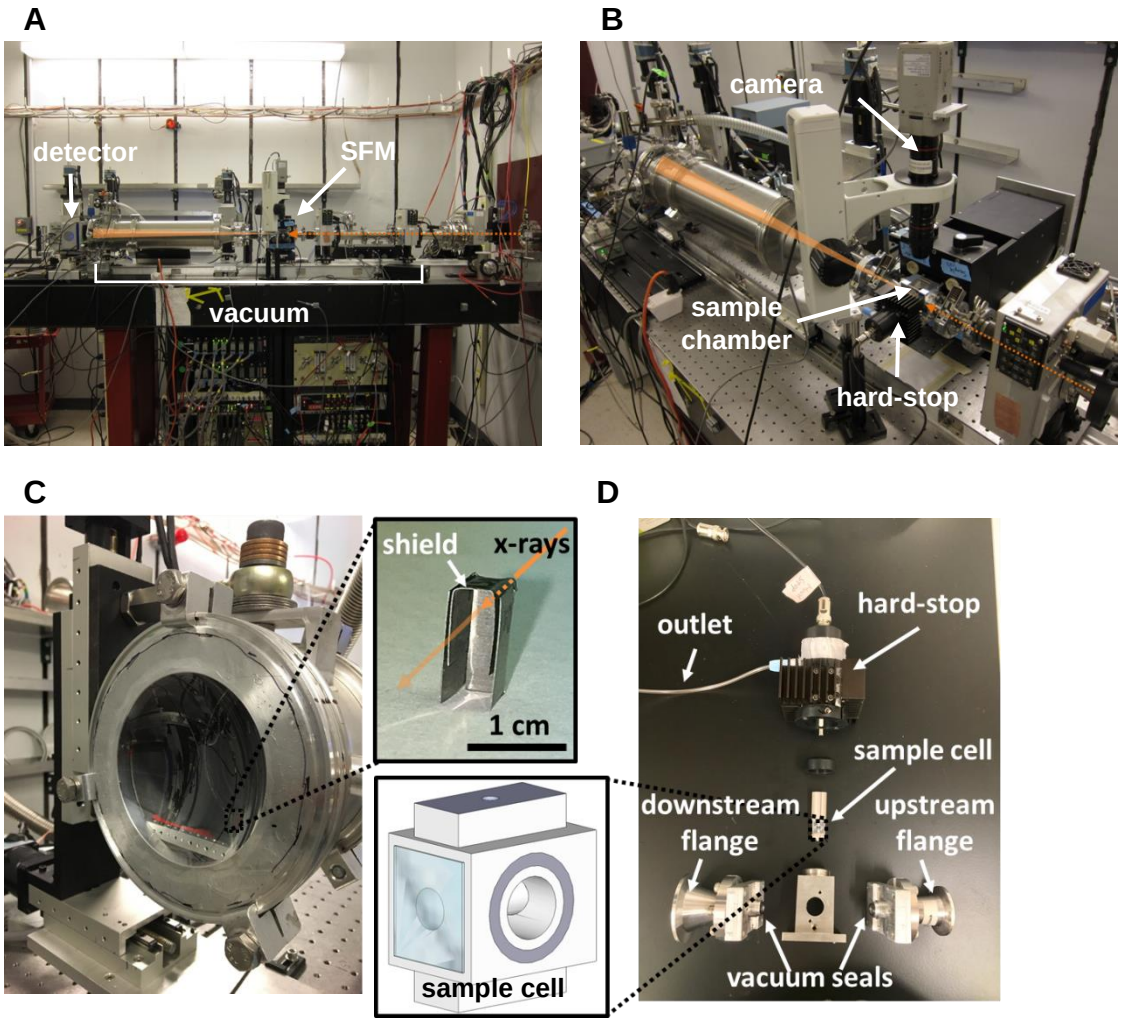


Figure 2.6 Images of SAXS beamline set up for time-resolved experiments at G1 station**,** Cornell High Energy Synchrotron Source, Cornell University, Ithaca, NY. (**A**) Side view of beamline. Incident x-rays (orange arrow) from the right are scattered by samples in the SFM (black box with blue tape). Scattered x-rays are imaged onto the detector on the left. Most of the chamber is maintained under vacuum to minimize parasitic scattering from air and windows. (**B**) View of stopped flow mixer mounted sideways into the beamline. The camera mounted on top provides a view into the sample chamber. (**C**) View of the end of the flight tube (detector moved away) showing where the beamstop is mounted. *Inset*: Example of a semi-transparent beamstop folded out of molybdenum foil. Incident x-rays from behind the beamstop are attenuated and imaged onto the detector. Folded shields block the diffraction ring from molybdenum. (**D**) Exploded view of SFM sample chamber. *Inset:* Cartoon drawing of the vacuum compatible sample cell. A vertically mounted quartz capillary holds liquid samples during x-ray exposure.

Figure 2.7 Images of SAXS beamline set up for time-resolved experiments at BioCAT Sector 18ID at Advanced Photon Source, Argonne National Lab, Lemont, IL. (**A**) Panoramic view of beamline setup for time-resolved experiments. (**B**) Closeup view of the SFM mixer. The vertical white syringes are used for loading samples and the camera providing a view into the observation chamber is mounted in front (lens tube visible). The upstream and downstream portions of the flight tube have mica windows and are brought as close to the sample as possible to minimize air gaps. (**C**) View of SFM mounted vertically onto an X-Z stage using unusually large C-clamps. The bottom and front panels of the SFM were removed to fit the mixer, exposing two of the stepper motors that drive the syringes at the bottom.

detector, each pixel can store up to 20 bits (up to 1,048,576 counts). Note that for the high flux used in time-resolved experiments, it is recommended to use thicker foil and stay within the lower region of this limit to reduce the chance of dropping counts (e.g. missing a photon when two arrive nearly simultaneously). The count rate limit for a pixel is 2×10^6 photons/s. When using molybdenum (Mo), the energy of the x-ray must be above 10 keV (at least 10.5 keV) to avoid any harmonic interactions with the Mo absorption K-edge (20 keV). To maximize the number of pixels available for collecting low-q scatter, the beam should be positioned as close to the edge of the foil as possible. The end of the foil should be folded towards the detector to form a shield (> 5 mm) to block diffraction peaks from molybdenum (Figure 2.6C). In order to maximize flux, a large beam (1 mm wide × 500 μm tall) can be used, but the beam should be focused at the beamstop for optimal scattering at low-q.

REFERENCES

[1] J.M. Tokuda, S.A. Pabit, L. Pollack, Protein-DNA and ion-DNA interactions revealed through contrast variation SAXS, Biophys. Rev. 8 (2016) 139–149.

[2] D.I. Svergun, M.H.J. Koch, Small-angle scattering studies of biological macromolecules in solution, Reports Prog. Phys. 66 (2003) 1735–1782.

[3] M.H. Koch, P. Vachette, D.I. Svergun, Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution., 2003.

[4] O. Glatter, O. Kratky, Small Angle X-ray Scattering Textbook (Glatter_Kratky_SAXS_1982).pdf, (1982).

[5] M.A. Graewert, D.I. Svergun, Impact and progress in small and wide angle X-ray scattering (SAXS and WAXS), Curr. Opin. Struct. Biol. 23 (2013) 748–754.

[6] S. Skou, R.E. Gillilan, N. Ando, Synchrotron-based small-angle X-ray scattering of proteins in solution., Nat. Protoc. 9 (2014) 1727–39.

[7] B.N. Chaudhuri, Emerging applications of small angle solution scattering in structural biology, Protein Sci. 24 (2015) 267–276.

[8] L. Pollack, Time resolved SAXS and RNA folding, Biopolymers. 95 (2011) 543–549.

[9] P. Bernadó, E. Mylonas, M. V Petoukhov, M. Blackledge, D.I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering., J. Am. Chem. Soc. 129 (2007) 5656–64.

[10] P. Bernadó, D.I. Svergun, Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering, Mol. BioSyst. 8 (2012) 151–167.

[11] G. Tria, H.D.T. Mertens, M. Kachala, D.I. Svergun, Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering, IUCrJ. 2 (2015) 207–217.

[12] A. Plumridge, S.P. Meisburger, L. Pollack, Visualizing single-stranded nucleic

acids in solution, Nucleic Acids Res. (2016) gkw1297.

[13] A. Plumridge, S.P. Meisburger, K. Andresen, L. Pollack, The impact of base stacking on the conformations and electrostatics of single-stranded DNA, Nucleic Acids Res. 45 (2017) 3932–3943.

[14] D. Orthaber, A. Bergmann, O. Glatter, SAXS experiments on absolute scale with Kratky systems using water as a secondary standard, J. Appl. Crystallogr. 33 (2000) 218–225.

[15] A. Guinier, La diffraction des rayons X aux tres petits angles: applications a l'etude de phenomenes ultramicroscopiques, Ann. Phys. (Paris). 12 (1939) 166–237.

[16] D.I. Svergun, M.H.J. Koch, Small-angle scattering studies of biological macromolecules in solution, Reports Prog. Phys. 66 (2003) 1735–1782.

[17] M.H.J. Koch, P. Vachette, D.I. Svergun, Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution, 2003.

[18] A. Guinier, G. Fournet, Small-Angle Scattering of X-Rays, Wiley, New York, 1955.

[19] J.P. Baldwin, P.G. Boseley, E.M. Bradbury, K. Ibel, The subunit structure of the eukaryotic chromosome, Nature. 256 (1975) 245–249.

[20] S. Bram, G. Butler-Browne, P. Baudy, K. Ibel, Quaternary structure of chromatin., Proc. Natl. Acad. Sci. 72 (1975) 1043–1045.

[21] J.F. Pardon, D.L. Worcester, J.C. Wooley, K. Tatchell, K.E. van Holde, Low-angle neutron scattering from chromatin subunit particles, J. Chem. Inf. Model. 2 (1975) 2163–2176.

[22] H.B. Stuhrmann, R.G. Kirste, Elimination Der Intrapartikularen Untergrundstreuung Bei Der Rontgenkleinwinkelstreuung an Kompakten Teilchen (Proteinen), Zeitschrift Fur Phys. Chemie-Frankfurt. 46 (1965) 247-.

[23] R.P. Hjelm, G.G. Kneale, P. Suau, J.P. Baldwin, E.M. Bradbury, K. Ibel, Small

angle neutron scattering studies of chromatin subunits in solution, Cell. 10 (1977) 139–151.

[24] D.I. Svergun, M.H.J. Koch, Structural Model of the 50 S Subunit of Escherichia coli Ribosomes from Solution Scattering, J. Mol. Biol. (1994) 66–77.

[25] D.I. Svergun, J.S. Pedersen, I.N. Serdyuk, M.H.J. Koch, Solution scattering from 50S ribosomal subunit resolves inconsistency between electron microscopic models, Proc Natl Acad Sci U S A. 91 (1994) 11826–11830.

[26] D.A. Jacques, J. Trewhella, Small-angle scattering for structural biology-Expanding the frontier while avoiding the pitfalls, Protein Sci. 19 (2010) 642–657.

[27] D.I. Svergun, M.H.J. Koch, P.A. Timmins, R.P. May, Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules, Oxford University Press, 2013.

[28] H.B. Stuhrmann, Neutron small-angle scattering of biological macromolecules in solution, J. Appl. Cryst. 7 (1974) 173–178.

[29] H.B. Stuhrmann, A. Miller, Small-Angle Scattering of Biological Structures, J. Appl. Crystallogr. 11 (1978) 325–345.

[30] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.

[31] D. Svergun, C. Barberato, M.H. Koch, CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 28 (1995) 768–773.

[32] R.P. Rambo, J. A. Tainer, Super-Resolution in Solution X-Ray Scattering and Its Applications to Structural Systems Biology, Annu. Rev. Biophys. 42 (2013) 415–441.

[33] C.D. Stoddard, R.K. Montange, S.P. Hennelly, R.P. Rambo, K.Y. Sanbonmatsu, R.T. Batey, Free State Conformational Sampling of the SAM-I Riboswitch Aptamer Domain, Structure. 18 (2010) 787–797.

[34] A. V Kazantsev, R.P. Rambo, S. Karimpour, J. Santalucia, J. a Tainer, N.R. Pace, Solution structure of RNase P RNA., RNA. 17 (2011) 1159–1171.

[35] B.L. Henke, E.M. Gullikson, J.C. Davis, X-Ray Interactions: Photoabsorption, Scattering, Transmission, and Reflection at E = 50-30,000 eV, Z = 1-92, At. Data Nucl. Data Tables. 54 (1993) 181–342.

[36] K. Luger, A.W. Mäder, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution., Nature. 389 (1997) 251–260.

Chapter 3

Revealing transient structures of nucleosomes as DNA unwinds †

Abstract

The modulation of DNA accessibility by nucleosomes is a fundamental mechanism of gene regulation in eukaryotes. The nucleosome core particle (NCP) consists of 147 bp of DNA wrapped around a symmetric octamer of histone proteins. The dynamics of DNA packaging and unpackaging from the NCP affect all DNA based chemistries, but depend on many factors, including DNA positioning sequence, histone variants and modifications. While the structure of the intact NCP has been studied by crystallography at atomic resolution, little is known about the structures of the partially unwrapped, transient intermediates relevant to nucleosome dynamics in processes such as transcription, DNA replication and repair. We apply a new experimental approach combining contrast variation with time-resolved small angle x-ray scattering (TR-SAXS) to determine transient structures of protein and DNA constituents of NCPs during salt-induced disassembly. We measure the structures of unwrapping DNA and monitor protein dissociation from *Xenopus laevis* histones reconstituted with two model

---

† **This chapter is reprinted with permission from** [1]**:** Y. Chen*, J.M. Tokuda*, T. Topping, J.L. Sutton, S.P. Meisburger, S. A. Pabit, L.M. Gloss, L. Pollack, Revealing transient structures of nucleosomes as DNA unwinds, Nucleic Acids Res. 42 (2014) 8767–8776. *These authors contributed equally

NCP positioning constructs: the Widom 601 sequence and the sea urchin 5S ribosomal gene. Both constructs reveal asymmetric release of DNA from disrupted histone cores, but display different patterns of protein dissociation. These kinetic intermediates may be biologically important substrates for gene regulation.

3.1 Introduction

Packaging of DNA by proteins in the nucleosome core particle (NCP) affects all DNA-based chemistries, including transcription, replication, repair and recombination [2]. The canonical NCP consists of 147 bp DNA wrapped around a symmetric histone octamer of two H2A-H2B heterodimers and an $(H3\text{-}H4)_2$ tetramer. Since DNA accessibility is a pre-requisite for initiating transcription and replication, an important unresolved question is how NCP disassembly proceeds to permit access to the DNA. Experiments carried out as a function of increasing salt (typically [NaCl] [3]) or force [4] reveal equilibrium intermediates with varying degrees of unwrapped DNA. Some contain bound, but disrupted protein cores [5–7]. However kinetic studies have been limited to FRET observations of spontaneous changes in DNA conformation [8–10] or to changes detected during protein binding [11,12]. The former measurements reveal that the dissociation of the DNA ends from the NCP (often termed breathing) occurs on a time scale of 100-250 ms, while larger scale openings involving the release of internal DNA segments occur on the order of 1-10 minutes.

No kinetic experiments to date have monitored both the DNA and histone components of the NCP as DNA is released and the octamer core is disrupted. Here we describe a novel approach that combines contrast variation with time-resolved small

angle x-ray scattering (TR-SAXS) to observe transient NCP structures following a rapid increase in salt. It is generally accepted that increasing concentrations of NaCl populate intermediates species similar to those transiently populated at low ionic strength [6,13] and this perturbation approach facilitates the application of alternative biophysical methods to characterize intermediate species and relative time scales for their formation. Two unresolved questions that may have major functional roles for chromatin *in vivo* are addressed: (1) Is DNA release from the histone core symmetric or asymmetric? (2) Do the eight histone proteins remain bound to DNA upon DNA unwrapping? Asymmetric disassembly of the nucleosome has been proposed based on the 5'-to-3' processivity of DNA and RNA polymerases, and the ability of RNA pol II to displace H2A-H2B dimers [14]. FRET-based models have generally assumed that the rapid DNA breathing motions are symmetrical [11], but transient NCP intermediates observed in recent high speed AFM measurements detected asymmetrical opening on similar time scales [15]. Equilibrium data indicate that the H2A-H2B dimers dissociate cooperatively from the NCP [7,16–18], but there is evidence for dissociation of one H2A-H2B dimer to form a hexasome kinetic intermediate [5,10,13,19].

We studied NCPs reconstituted with *Xenopus laevis* histones and two well-characterized 149 bp NCP-positioning sequences: the high affinity 601 DNA developed by the Widom lab (601-NCP) [7,20] and the weaker positioning "5S DNA" from the promoter region of the *Lytechinus variegatus* (sea urchin) 5S ribosomal gene (5S-NCP) [21].

Small angle x-ray scattering (SAXS) is a label free technique that reports the global conformation and composition of macromolecules in solution [22–27]. The scattered

intensity provides information about the average composition, size, and shape of the scattering particles. The extrapolated scattering intensity at zero angle, I(0), is proportional to the square of the excess electron density of the particles in solution and is therefore sensitive to changes in the oligomeric state of the complexes. Thus I(0) can be used to monitor the dissociation of proteins from the NCP. A quantitative measure of size is reported as the radius of gyration (Rg). For scatterers with homogenous electron densities, the scattered intensity I(q) is directly related to macromolecular shape. However, for complexes with components that have varying electron densities (e.g. protein and nucleic acids), the relationship between I(q) and macromolecular shape becomes ambiguous. The simplest way to circumvent this challenge is to apply contrast variation and match the electron density of solvent with the lower density protein (See Appendix A S.I. Contrast Variation). By adding 50% sucrose to the solvent, the protein becomes invisible above the background and only the DNA contributes to the scattering (Figure 3.1). Contrast variation SAXS has successfully revealed the structure of RNA or DNA complexed with proteins in static studies [28,29]. Here we describe the application of contrast variation to monitor changing NCP conformations as [NaCl] is increased in equilibrium titrations. We then expand on this strategy by incorporating a stopped-flow mixer (SFM) to measure time dependent changes following the rapid addition of salt (Figure 3.2). Extensive characterization of mixing performance verified a $\approx$ 5 ms mixing dead time, even for viscous sucrose solutions (See Appendix A S.I. Mixer Characterization).

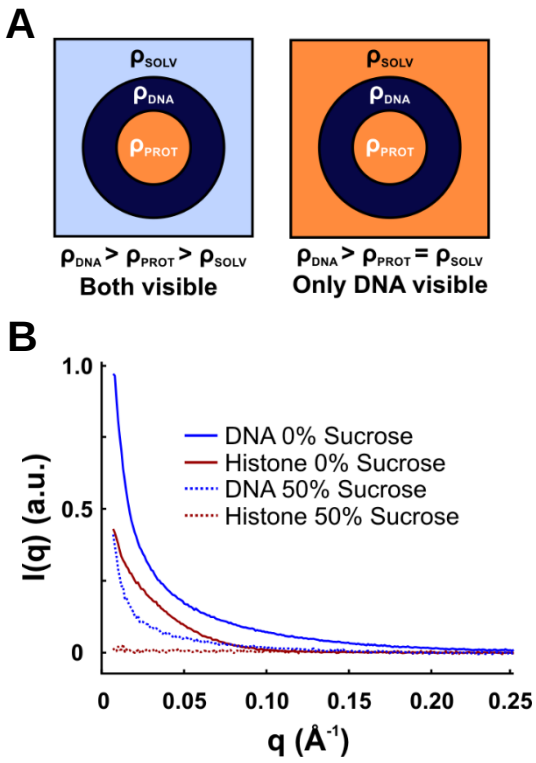


Figure 3.1 Contrast variation reveals DNA conformation within protein-nucleic acid complexes. (**A**) Cartoon depicting how contrast variation is used to isolate scattering from the DNA component in protein-nucleic acid complexes. Left: The protein-nucleic acid complex in solution can be approximated as three phases with electron densities ρ_{SOLV} (light blue), ρ_{PROT} (orange), and ρ_{DNA} (dark blue). Right: Since contrast arises from electron density differences, the electron density of the solvent is increased by adding small molecules such as sucrose until it matches that of the protein. Consequently, the protein is effectively "blanked" and only the DNA contributes to the measured scattering signal. (**B**) Scattering profiles for NCP components measured separately in 2 M NaCl with and without sucrose. In 50% sucrose, proteins become invisible above the background and only the DNA contributes to the scattering. The resulting signal for the DNA is decreased because of the reduced contrast between the DNA and solvent.

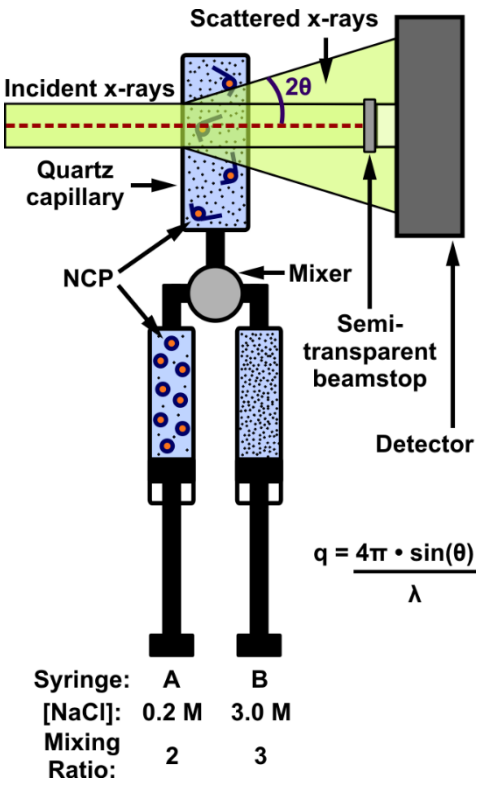


Figure 3.2 Schematic of stopped-flow mixing experiment to probe salt-induced disassembly of NCPs without sucrose. Compact NCPs in 0.2 M NaCl mix with buffer containing 3.0 M NaCl to achieve a final NaCl concentration of 1.9 M, where full NCP disassembly occurs. The optimal flow rates and volumes used were 6 mL/s and 315 µL for 0% sucrose and 7.5 mL/s and 375 µL for 50% sucrose. In 0% sucrose, both nucleosomal DNA and histones are "visible," hence TR-SAXS data reports changes in NCP global size, structure and composition. In 50% sucrose, only nucleosomal DNA is "visible," TR-SAXS data directly reveals changes in DNA conformation. λ is the wavelength of the incident x-rays in Angstroms.

3.2 Materials and Methods

3.2.1 NCP production and reconstitution

The standard procedures used to express and purify *Xenopus laevis* histones are described in Appendix A S.I. Histone Purification. The production of both DNA sequences (Widom 601 and the 5S sequences) is also described in Appendix A S.I. DNA Production.

Our previous NCP reconstitutions employed size-exclusion HPLC in 2 M KCl to purify histone octamers before the addition of DNA and step-wise dialysis to lower salt concentrations [7]. If care is taken to optimize the dimer:tetramer:DNA ratios, NCP preparations of similarly high homogeneity can be obtained without initial purification of the octamer [30]. Histone octamers were formed by mixing H2A-H2B and $(H3\text{-}H4)_2$ in 2 M NaCl, 0.1 mM EDTA and 20 mM TrisCl pH 7.5 and incubating on ice overnight before the addition of DNA. Using the button-dialysis method [31], preliminary small-scale reconstitutions were done by step dialysis (2 M to 0.85, 0.65, 0.2 and 0 M NaCl) to determine the optimal conditions for every preparation, varying both the H2A-H2B:$(H3\text{-}H4)_2$ and octamer:DNA ratios. The quality of the NCP samples were analyzed by 5% native gel PAGE (described in detail in ref. [30]) to identify the ratio that resulted in NCPs with no free DNA or other histone-DNA complexes. Subsequent large-scale NCP preparations combined the appropriate amount of histone oligomers and DNA to a final concentration of 5 µM NCP, followed by step dialysis. Samples with the 5S DNA were heat-shifted by incubation at 37°C for 1 hour to achieve homogeneously positioned NCPs. NCP samples were concentrated to ≈ 50 µM by centrifugal ultrafiltration.

3.2.2 Equilibrium SAXS experiments

SAXS data were collected at the Cornell High Energy Synchrotron Source (CHESS) G1 station with an x-ray energy of 10.5–10.6 keV. Sample-to-detector distance was measured to be ≈ 1 m using a silver behenate standard. The available q-range was from ≈ 0.007–0.25 $Å^{-1}$. Samples and matching buffers were manually prepared and equilibrated for at least 5 minutes before being loaded into a 2 mm diameter quartz capillary with 10 μm walls (HR6-150, Hampton Research). Samples were oscillated during x-ray exposure to reduce radiation damage. SAXS profiles of matching buffers were measured before and after each sample to monitor beam conditions and ensure capillaries were clean. The scattered x-rays were imaged onto a photon counting array detector (PILATUS 100K, Dectris). Multiple images with 1–30 s exposures were acquired for each sample, and exposure dependent changes reflecting radiation damage were carefully monitored. Incident beam was measured for normalization using either a PIN diode embedded in the beamstop or a semi-transparent beamstop to directly image the attenuated beam. An NCP concentration series in 0.2 M NaCl showed negligible concentration dependent inter-particle interactions at ≈ 5–10 μM. Thus equilibrium and time-resolved experiments were conducted at these concentrations.

3.2.3 Time-resolved SAXS experiments

Time resolution was achieved by incorporating a stopped flow mixer (SFM-400, Bio-Logic) as shown in Figure 3.2. Custom sample cells with 2 mm path length quartz capillaries were used, allowing for direct comparison with equilibrium experiments after adjusting for contrast. NCPs in low salt (0.2 M) were mixed with high salt buffers (3

M) at a ratio of 2:3 to achieve a final NaCl concentration of 1.88 M. Efficient mixing of viscous solutions in the contrast matched condition was ensured by (1) loading both the NCP sample and high salt buffer with 50% sucrose (so the mixing solutions have similar viscosities), (2) incorporating a high-density mixer (model HDS, Bio-Logic), and (3) optimizing the mixing protocol (See Appendix A S.I. Mixer Characterization). Matching buffers were measured before and after each experiment by replacing the sample syringe with low salt buffer.

Several design features were incorporated to maximize the measured SAXS signals and ensure reliable data collection. Background scattering from air and windows was minimized by placing the sample capillary and x-ray flight path under vacuum. A semi-transparent molybdenum beamstop was used to attenuate and image the beam for reliable normalizations. The PILATUS 100K detector was operated in "movie mode" with a time frame of 20 ms (17 ms exposure + 3 ms readout). Since the samples were not oscillated after mixing, the NCPs were susceptible to radiation damage. For samples without sucrose, attenuators were placed in the beam after 10 s to limit radiation damage. Interestingly, samples with 50% sucrose appeared to be less susceptible to radiation damage, thus no attenuators were used.

3.2.4 Data analysis

All SAXS images were processed using MATLAB (MathWorks). SAXS intensity patterns for each image were azimuthally averaged about the beam center and SAXS profiles from multiple images of the same sample were averaged to improve the signal-to-noise ratio. Uncertainties in I(q) for each image were estimated as standard deviations

divided by the square root of the number of pixels binned for each q-value and propagated appropriately. SAXS curves for NCPs were determined by subtracting the scattering curves of matching buffers from the total scattering curves of the samples. Raw data are displayed as Kratky plots, I(q)•q2 vs q. More globular objects display strong peaks in Kratky plots. Pair-wise distance distributions, P(R), were calculated using the regularized indirect Fourier transform program GNOM [32]. High-q values were omitted in the fits when they largely affected P(R) shape in order to avoid artifacts from including data with low signal-to-noise ratios. The largest dimension of the molecule (Dmax) was systematically varied until (1) a good fit to the data was achieved, (2) the P(R) shape was stabilized, and (3) the P(R) had a smooth decaying tail. All I(0)s and Rgs reported were calculated using GNOM [32].

For time-resolved studies, the SAXS profiles from the 20 ms exposures were binned to improve signal-to-noise ratio (at the expense of time resolution). Optimal bin sizes were determined from kinetic SVD analysis [33] (Supplementary Figure A.4 and A.5). SAXS profiles for corresponding time bins from subsequent experiments (4-6 repeats for each condition) were averaged. I(0,t), Rg(t) and P(R,t) analysis and modeling for each time point (t) were conducted following the same strategies as the equilibrium experiments. In 50% sucrose, SAXS curves extrapolated to nearly identical I(q = 0,t)s, indicating only DNA contributed to the SAXS signal.

3.2.5 Modeling and P(R) analysis

Model structures were utilized to gain physical insight into the features observed in P(R) functions. The DNA component of the crystal structure for the nucleosome core particle

(1AOI) was used to model the NCPs in the completely wrapped state. A linear 149 bp DNA using the Widom 601 DNA sequence was generated using Nucleic Acid Builder [34] to model the NCPs in the free unwrapped state. Minor differences in length or sequence identity did not make significant differences. To account for solvation, theoretical scattering profiles were first calculated from the atomic coordinates of the models using the program CRYSOL [35]. Theoretical scattering profiles were then processed through GNOM to determine P(R) and Rg. As shown in Figure 3.4D, structural features from three length scales were identified. Alternate conformations with the DNA released by varying degrees were generated by appending linear DNA fragments to the DNA component of 1AOI (Supplementary Figure A.2). The trends observed in P(R) as DNA was unwrapped further validate the proposed interpretation of the features (Supplementary Figure A.2).

3.3 Results

3.3.1 Equilibrium SAXS reveals salt induced NCP disassembly

As a precursor to time resolved studies, we first measured the equilibrium response of the NCP to increasing [NaCl], both with and without sucrose (Figure 3.3 and Supplementary Figure A.1). In buffer containing 200 mM NaCl, NCPs are compact with DNA predominantly wrapped around the histone core. At lower ionic strength, repulsive forces are detected between NCPs in the concentration range of interest, leading to interparticle interference effects that distort the SAXS profiles at the lowest angles (as observed in ref. [36]). Increasing concentrations of NaCl (from 0.2 M to

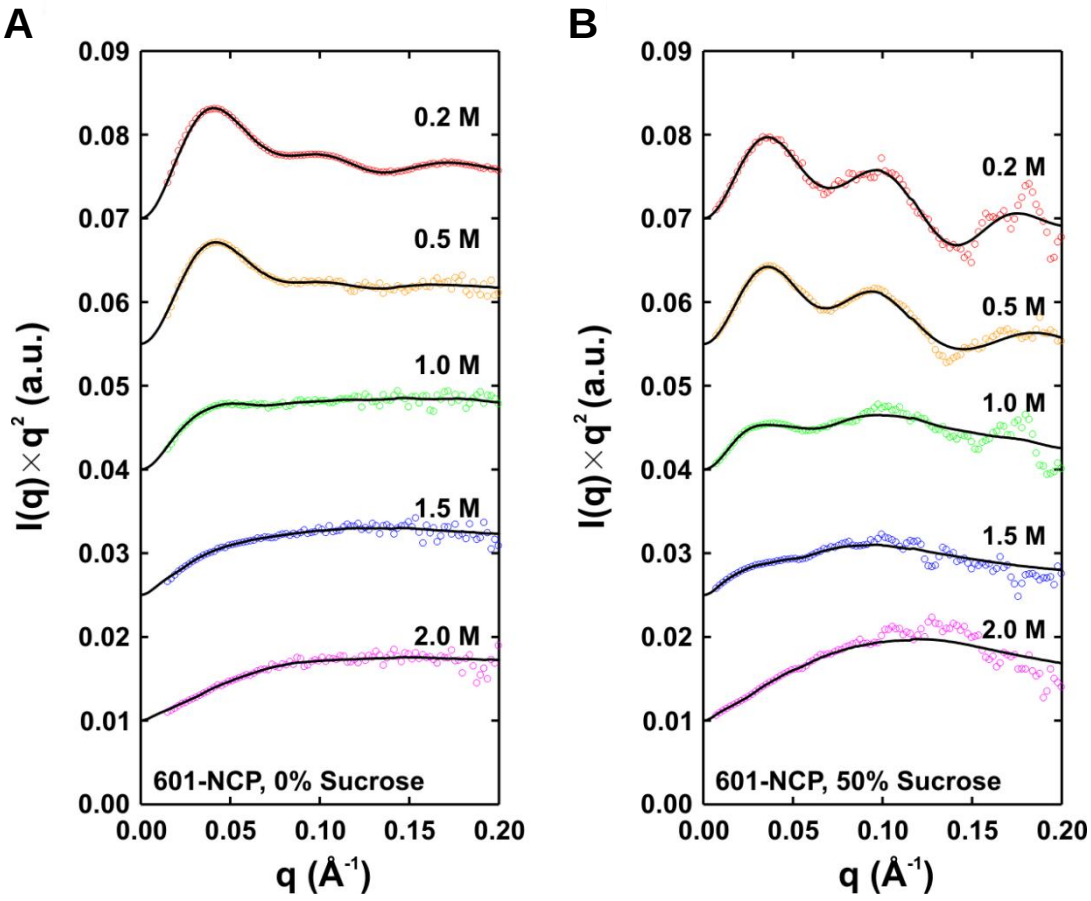


Figure 3.3 Kratky plots for 601-NCP in varied [NaCl] with (**A**) 0% and (**B**) 50% sucrose. The data (colored circles) and regularized fits to the data (black lines) are scaled and offset to enhance visualization. Since the data in (**B**) is significantly noisier, a moving average with a span of 9 was used to show the quality of the fits. The transition from a compact to an extended structure is observed as the strongly peaked curve changes to a more plateaued curve at high q with increasing [NaCl].

2 M) populate intermediate species similar to those transiently observed at physiological ionic strength (0.05-0.1 M monovalents) [6,13]. Depending on the DNA sequence, dissociation of H2A-H2B dimers and the $(H3\text{-}H4)_2$ tetramer occur above ≈ 0.6 and ≈ 1.5 M NaCl, respectively [7,18]. At 2.0 M NaCl, the histone proteins are largely dissociated from the DNA. Importantly, FRET based studies [7] showed that sucrose

has no measurable effects on the NaCl-dependent equilibrium stability of NCPs.

The [NaCl]-dependent SAXS profiles measured with and without sucrose for the 601- and 5S-NCP constructs show a response to NaCl similar to that reported in previous biophysical assays [6,7,13,18]. Raw data, shown as Kratky plots for 601-NCP (Figure 3.3) and 5S-NCP (Supplementary Figure A.1) reveal dramatic conformational transitions from globular to extended structures as the NaCl concentration increases from 0.2 M to 2.0 M. The structural details (observed as peaks and troughs in the curves) become significantly more pronounced when the proteins are blanked in sucrose, highlighting the power of contrast variation (Figure 3.3B and Supplementary Figure 1B). The extension of the NCP with increasing [NaCl] is also reflected by salt dependent Rgs. (Figure 3.4A). These raw SAXS data reveal subtle differences between the two constructs. Although the general size for the two constructs is comparable in the compact state at 0.2 M NaCl (43.7 ± 0.5 Å for 601-NCP and 44.9 ± 0.3 Å for 5S-NCP), the broadened peaks and troughs in the Kratky profiles in 50% sucrose suggest a less well-defined structure for the 5S-NCP DNA compared to the 601-NCP DNA (Figure 3.3 and Supplementary Figure A.1). Furthermore, at 1 M NaCl, Rg values show that the 5S-NCP DNA is more extended than 601-NCP DNA (Figure 3.4A) suggesting that it is more readily unwrapped by increasing [NaCl]. These differences likely reflect weaker DNA-histone interactions for the 5S sequence relative to 601 [20].

3.3.2 DNA conformation revealed by P(R) analysis and ensemble optimization method

The dramatic conformational changes, already revealed in the Kratky plots of

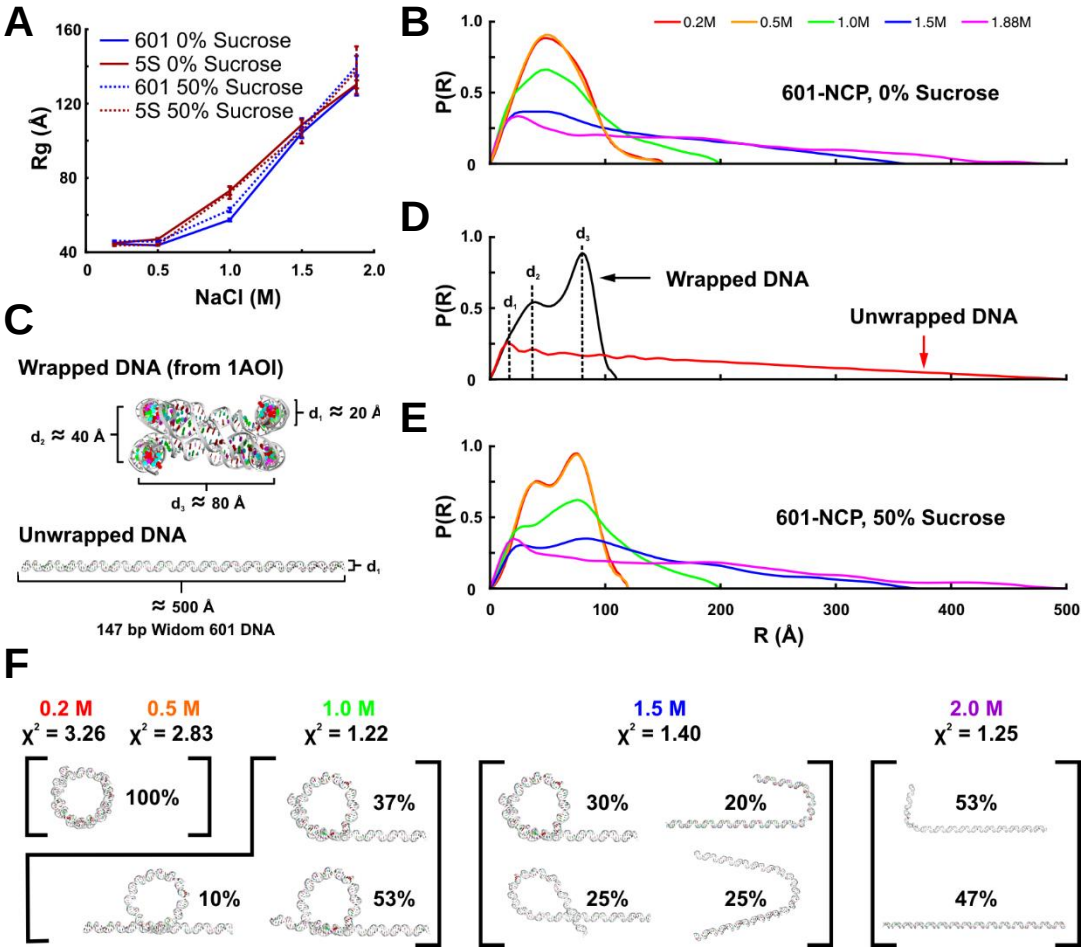


Figure 3.4 Contrast variation reveals DNA conformation during salt-induced disassembly. (**A**) Radius of gyration (Rg) for 601-NCP and 5S-NCP in equilibrium with different NaCl concentration with and without sucrose. An expansion in size from 45 Å to 130 Å is detected for both constructs with increasing [NaCl]. At 1 M NaCl and 50% sucrose, the 5S-NCP DNA (72 ± 3 Å) is more expanded than the 601-NCP (63 ± 1 Å). (**B**) P(R)s for 601-NCP in equilibrium with varied [NaCl] and 0% sucrose. A general extension of the NCP is observed with increasing [NaCl]. (**C**) Models used for calculating theoretical P(R)s for the wrapped (DNA component from the NCP crystal structure 1AOI) and unwrapped (linear 147 bp Widom 601 DNA) states. (**D**) P(R) peaks at three length scales are attributed to structural features as follows: d_1 – diameter of duplex DNA; d_2 – distance between overlapping DNA ends; d_3 – diameter of wrapped structure. (**E**) P(R)s for 601-NCP in equilibrium with varied [NaCl] and 50% sucrose. With the proteins blanked, P(R) reveals how DNA conformation changes as the NCP is destabilized by increasing [NaCl]. (**F**) Models representing ensembles of conformations selected by EOM that produces theoretical SAXS profiles that best fit the [NaCl]-dependent SAXS data. Percentages reporting weights of models and χ^2 values assessing overall fit to experimental SAXS data are shown.

Figure 3.3, are most readily interpreted by examining pairwise distance distribution functions (P(R)s) computed from SAXS profiles using GNOM [32]. Real-space information is displayed in P(R)s as histograms of all intra-molecular distances R (Figure 3.4B,D,E). Peaks in P(R) represent length scales that are repeated within the particles. The distribution function approaches zero at the largest intra-molecular dimension (Dmax). In the absence of sucrose, P(R) is challenging to interpret because the contributions from protein and DNA components cannot be distinguished. These P(R)s reveal only general features, such as the overall size and largest dimension of the NCPs (Figure 3.4B). At low [NaCl], compact NCPs are characterized by P(R) curves with prominent peaks near 50 Å and a full-width of approximately 100 Å, consistent with the largest dimension of wrapped NCPs. Similar P(R) curves for full NCPs (in 0% sucrose and 0.2 M NaCl) have been reported for several different DNA sequences [23–27]. With increasing [NaCl], NCPs are destabilized, and DNA unwrapping is observed as the general extension of the P(R) curves to approximately 500 Å, the length of the free dsDNA in solution.

When sucrose is added and the proteins "disappear," clear and identifiable features emerge in the [NaCl]-dependent P(R) curves, revealing specific conformational details of the unwrapping DNA (Figure 3.4E, Detailed in Supplementary Figure A.2). To interpret the features present in these histograms, we computed the scattering profiles of model structures using the wrapped and unwrapped models of nucleosomal DNA (Figure 3.4C-D). As the DNA dissociates: (1) a peak appears at $d_1 \approx 20$ Å, revealing an extension of DNA duplex, (2) the peak at $d_2 \approx 40$ Å disappears, corresponding to the decrease in overlap between DNA ends and (3) the peak at $d_3 \approx 80$ Å decreases,

corresponding to a disruption of the overall wrapped structure. Clearly, contrast variation SAXS provides incisive structural information about DNA conformations accompanying [NaCl]-induced NCP dissociation.

To elucidate the configuration of nucleosome-bound DNA at different NaCl concentrations, we generated a pool of candidate DNA conformations for comparison with the contrast matched data. This pool contains both symmetric and asymmetric structures (representative structures from the pools are shown separately in Supplementary Figure A.8). Because multiple states may be present at intermediate NaCl concentrations, the ensemble optimization method (EOM) [37] was used to identify collections of structural models that best recapitulate the [NaCl]-dependent SAXS data (See Appendix A S.I. Ensemble Optimization Method). The DNA models shown in Figure 3.4F were selected by EOM analysis from a pool of 32 models, containing both symmetric and asymmetric structures. In 0.2-0.5 M NaCl, the ensemble is relatively monodisperse, with DNA wrapped around the histone core. As the NaCl concentration increases to 1.0 M, roughly half of the DNA is released, but the DNA ends are still crossed. At 1.5 M NaCl, more of the DNA is released but an increased heterogeneity appears in the chosen models. Finally at 2.0 M NaCl, most of the DNA is released. In the ensembles that represent 1 M and 1.5 M NaCl, more than 75% of the models selected are asymmetric. The relationship between peaks in P(R) at the three length scales d_1, d_2, d_3 and the structural features of the ensembles are consistent with interpretations based on P(R) curves computed for the representative model structures (Supplementary Figure A.2).

3.3.3 Time-resolved SAXS reports changes in NCP composition and size during disassembly

To measure the structures of transient states accompanying [NaCl]-dependent NCP dissociation, we coupled SAXS with a stopped flow mixer (Figure 3.2). Fully wrapped NCPs in 0.2 M NaCl were rapidly mixed with buffer containing 3.0 M NaCl in a 2:3 ratio to achieve a final NaCl concentration of 1.9 M. Time-resolved SAXS measurements were carried out with and without sucrose. Representative curves at different time points after mixing are shown in Supplementary Figure 3A-B. At 0% sucrose, the time-dependent dissociation of histone components from NCPs is revealed by monitoring changes in I(0) [22]. I(0) is related to the molecular mass of NCP, but interpretation of I(0) requires knowledge of both sample heterogeneity and contrast. It is important to note that the contrast depends not only on sucrose, but also on [NaCl]. Since a wide range of [NaCl] was used for the equilibrium experiments, we limited our analysis to the endpoints with the assumption of monodispersity (fully associated octamers in 0.2 M NaCl and fully dissociated in 2.0 M NaCl). Since time dependent changes using a stopped flow mixer are measured against a fixed background, the contrast does not change with time. Relative changes in I(0,t) reveals details of NCP dissociation.

Figure 3.5A-B shows I(0,t) for 601- and 5S-NCPs in 1.9 M NaCl without sucrose. Equilibrium values for the intact and completely disassembled NCPs are shown for

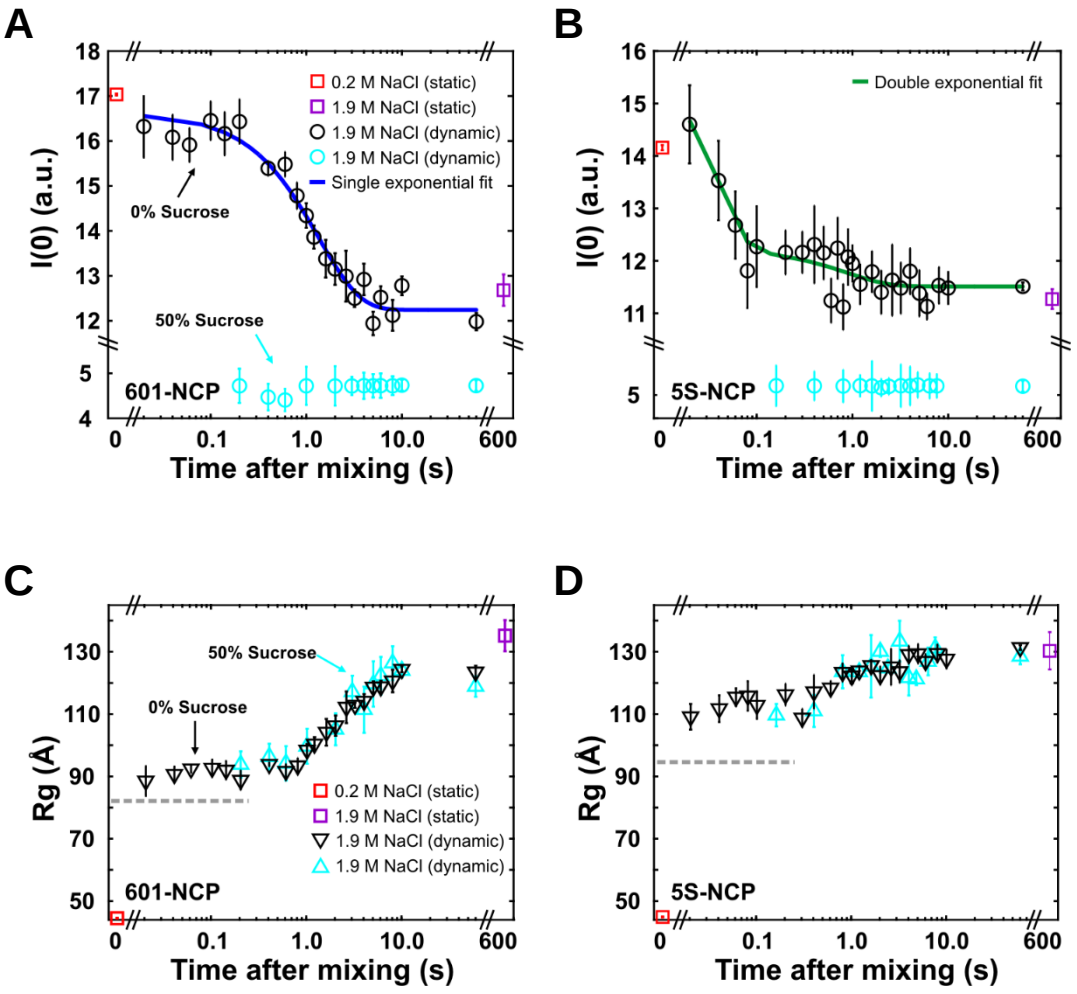


Figure 3.5 I(0,t) and Rg(t) analysis monitoring protein dissociation and expansion of NCP size as DNA is released**.** (**A**,**B**) In 0% sucrose, I(0) values (black circles) monitor the time dependent release of histone components for the 601-NCP and 5S-NCP, respectively. The 601-NCP remains intact for the first 200 ms and is described by a single exponential decay, whereas the 5S-NCP appears to dissociate faster, but shows a double exponential decay. In 50% sucrose, I(0) values (cyan circles) are decreased due to the reduced contrast and remain relatively unchanged since the signal arises from DNA alone. (**C**,**D**) Time dependent changes in the radius of gyration reveals DNA unwrapping for 601-NCP and 5S-NCP, respectively. Dynamics were monitored on time scales ranging from 20 ms to 60 s after mixing. Equilibrium values for intact (in 0.2 M NaCl, corrected for contrast – See Appendix A S.I. I(0) Analysis) and fully dissociated (in 2.0 M NaCl for 10 min) NCP states are shown for comparison. The gray dashed lines in (**C**) and (**D**) represent the R_gs for 601-NCP (= 82.15 Å) and 5S-NCP (= 95.37 Å) if the 'J' shaped DNAs (Figure 3.6C-D) are bound to intact histone octamers (models shown in Supplementary Figure A.7A).

comparison. The I(0) values have not been concentration corrected, so are arbitrarily scaled for each construct. For both sequences, I(0) shows no decrease at the earliest time point measured (20 ms), suggesting that all histone proteins remain bound to nucleosomal DNAs on this rapid time scale. For the 601-NCP, I(0) remains constant for the first 200 ms after mixing. Protein dissociation occurs with apparent first-order kinetics and a time constant of $0.74 \pm 0.08\ s^{-1}$. Significant variation is seen for 5S-NCPs where the time course of I(0) is well-described by a double exponential decay with two distinct rates: $41.6 \pm 13.9\ s^{-1}$ and $1.13 \pm 0.74\ s^{-1}$. Because the amplitude of the second phase is small, we also analyzed the time-binned SAXS curves using singular value decomposition, which projects each curve onto basis states and provides a sensitive measure of subtle changes (Supplementary Figure A.5C, see Appendix A S.I. Singular Value Decomposition). Two transitions emerge from this analysis with rates that are similar to those of I(0). Whereas the histone proteins dissociate in a single phase from 601 DNA, the two phase curve of the 5S construct indicates a different pattern of protein dissociation.

To assess the overall extent of the NCP structure, we computed the time dependent radius of gyration, Rg (Figure 3.5C-D), for both constructs and compared the values with those obtained from equilibrium studies for intact (≈ 45 Å) and fully disassembled (≈ 130 Å) NCPs. At the earliest measured time point, the Rgs of 90 Å for 601-NCP and 110 Å for 5S-NCP are larger than at 0.2 M NaCl, suggesting a significant expansion within the first 20 ms. Interestingly, for the 601-NCP, plateaus observed in both I(0,t) and Rg(t) suggest a kinetic intermediate that persists for the first ≈ 200 ms. The unchanging I(0) value indicates that all of the histone proteins remain bound, despite

expansion to a structure with Dmax ≈ 325 Å. In contrast, the rapidly decreasing I(0) values for 5S-NCP reveal that the protein core breaks up and dissociates in less than 0.1 s. By 10 s, both constructs reach a highly expanded size (≈ 500 Å), with minimal further increase.

3.2.4 NCP intermediate structures revealed by contrast variation TR-SAXS

In order to integrate structural models from time resolved data acquired in 0% and 50% sucrose, we first confirmed that sucrose has a minimal effect on the dissociation dynamics. SVD analysis of time-binned SAXS curves with and without sucrose (representative curves are shown in Supplementary Figure A.3) showed similar rates, suggesting that sucrose minimally alters NCP dissociation dynamics (Supplementary Figure A.4 and A.5). Furthermore, time dependent changes in Rg are not affected by the viscosity of the sucrose, suggesting that DNA unwrapping is not diffusion driven (Supplementary Figure A.7A).

Applying the P(R) analysis to the kinetic scattering profiles, we characterized the dynamic conformational changes of both constructs at 0% and 50% sucrose. In 0% sucrose, both P(R) distributions at t = 20 ms contain a peak similar to that observed in the fully bound NCP, but with Dmax values of 320 Å for 601-NCP and 385 Å for 5S-NCP (Figure 3.6A-B). At longer times, the major peak position shifts from ≈ 50 Å to ≈ 20 Å, characteristic of DNA release observed in static experiments (Figure 3.4B). These changes occur more rapidly for the 5S-NCP (Figure 3.6A-B and Supplementary Figure 6A-B) than for the 601-NCP, indicating its decreased stability.

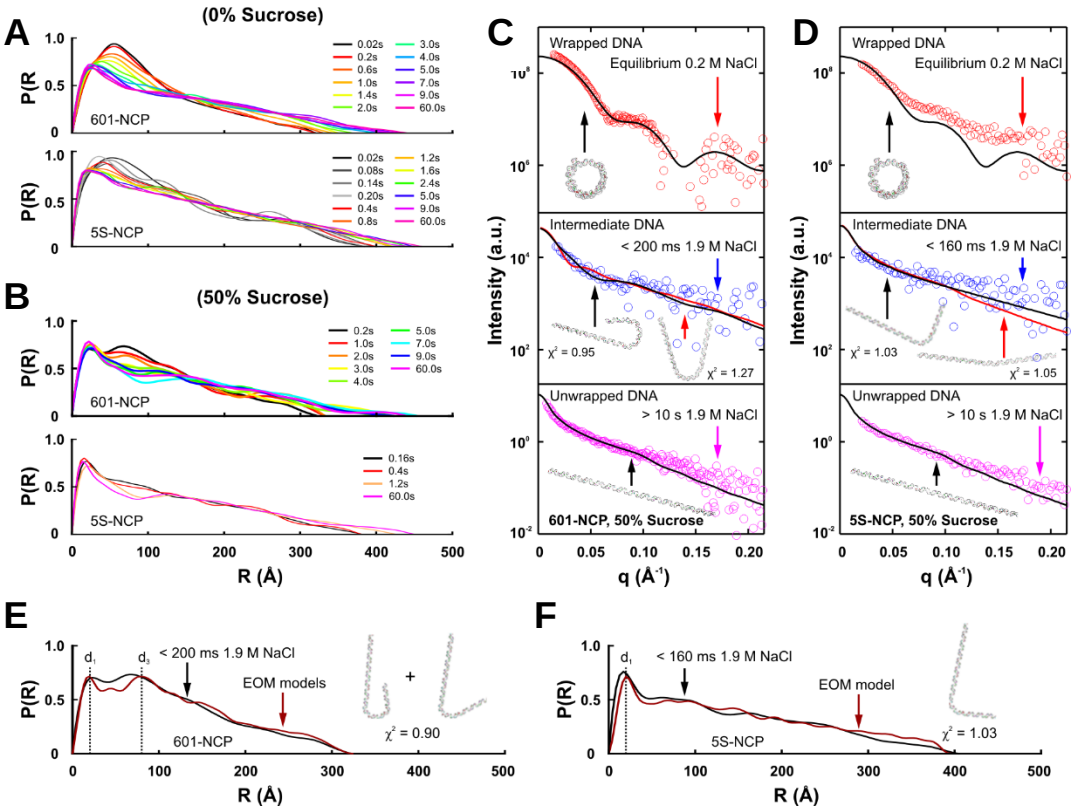


Figure 3.6 Time-resolved SAXS with contrast variation reveals DNA conformation of kinetic intermediates. (**A**) Pairwise distance distribution functions, P(R), computed from time-resolved scattering profiles of 601-NCP and 5S-NCP in 0% sucrose. (**B**) Time-resolved P(R)s for 601-NCP and 5S-NCP in 50% sucrose revealing DNA conformational changes during salt-induced disassembly. (**C**,**D**) Comparison of experimental scattering profiles for (**C**) 601-NCP and (**D**) 5S-NCP in 50% sucrose with best fitting theoretical scattering profiles for symmetric (black lines) and asymmetric (red lines) models for the wrapped, intermediate and unwrapped DNAs (offset to aid visualization). Theoretical profiles are calculated from the models shown as insets. The intermediate DNA models were determined using EOM and the goodness of fits was assessed by comparing χ^2 values. (**E**,**F**) P(R)s for the ensembles (red) selected by EOM analysis (models shown with χ^2 fit to SAXS data) compared with the experimental P(R)s (black) determined from (**E**) the 200 ms kinetic intermediate of the 601-NCP in 50% sucrose and (**F**) 160 ms data of the 5S-NCP in 50% sucrose. (For details see Appendix A S.I. Minimum chi-square (χ^2) fit, Appendix A Ensemble Optimization Method, and Supplementary Figure A.8).

In 50% sucrose, P(R) analysis provides structural information about the DNA conformation in the 601-NCP intermediate (t < 200 ms). Two characteristic peaks appear, at distances d_1 and d_3 (shown in Figure 3.6E). Comparison with static analysis of Figure 3.4D and Supplementary Figure A.2 suggests that in this state, the DNA has a very large dimension (Dmax ≈ 325 Å), does not overlap itself (missing d_2 peak at ≈ 40 Å) but nevertheless forms nearly a complete wrap around the histones (a pronounced d_3 peak at ≈ 80 Å). Based on geometrical arguments, these structural signatures support an asymmetric DNA conformation. A comparison of the experimental data with computations based on symmetric and asymmetric models from a library of candidate DNA structures showed strongest agreement for a 'J'-shaped asymmetric model. Goodness of fit was assessed by a $\chi 2$ test and ensemble optimization method (Supplementary Figure A.8). SAXS profiles and models of DNA release from 601-NCP and 5S-NCP are shown in Figure 3.6C-D. A similarly detailed kinetic analysis of 5S-NCP dissociation proved challenging due to its more rapid dissociation: the 5S-NCP is almost completely unwrapped on the millisecond timescale. However, P(R) and EOM analysis of the first 160 ms time-resolved SAXS data of 5S-NCP also support asymmetric release (Figure 3.6F and Supplementary Figure A.8).

Structural models of the 601-NCP kinetic intermediate can be refined by integrating all of the above data. The DNA conformation is asymmetric, all proteins remain associated and the Rg measured for the 601-NCP intermediate (≈ 90 Å, Figure 3.5C) is significantly larger than that computed for models of the NCP with the histone intact as an octamer on the wrapped end (≈ 82 Å, Supplementary Figure A.7A). With the DNA constrained to a ‘J’-shape, this dramatic increase in Rg is best explained by NCP models

where protein-protein interactions are disrupted but protein-DNA interactions are not. All histone proteins remain bound, but the octamer core is no longer integral. Additional support for this model is provided from full scattering profiles (Supplementary Figure A.7B).

In contrast, the 5S-NCP exhibits much faster salt-induced dissociation dynamics, with no stable intermediate detected. Despite these differences, it is notable that the 5S DNA appears to unwrap asymmetrically, like the 601 DNA. This assessment is based on P(R) and EOM analyses of the first 160 ms TR-SAXS curve with 50% sucrose (Figure 3.6D,F) and suggests that asymmetric release is a common feature among different DNA sequences. The histone core is disrupted before dissociation, based on the Rg analysis (Figure 3.5D and Supplementary Figure A.7A).

3.4 Discussion

The present study reveals the conformations of nucleosomal DNA in response to increasing concentrations of NaCl. NCP disassembly induced by increasing [NaCl] proceeds through steps that mimic those observed in the nucleus [6,13]. Thus, it is reasonable to hypothesize that the conformational changes of the NCP accessed by varying salt concentrations probe the inherent dynamic properties of the NCP that dictate interactions with cellular machinery during chromatin function.

Representative DNA models that best describe the SAXS data for 601-NCPs in equilibrium with varying [NaCl] are shown in Figure 3.4F. Interestingly, many of the models selected contain asymmetrically extended DNA. This type of release may reflect the asymmetric affinity of the Widom 601 sequence reported by Chua et al [38].

Detailed analysis of the 5S-NCP was less straightforward, possibly due to heterogeneity that arises from multiple translational settings. The application of contrast variation in conjunction with modelling reveals interesting differences between the former sequence, which was engineering for strong binding, and the latter, weaker positioning sequence. Future work with other sequences has the potential to reveal additional dissociation pathways and may elucidate the complex, but very important connections between sequence, spacing and affinity.

Figure 3.7 presents a schematic timeline for the salt-induced disassembly of the 601-NCP, as well as the less stable 5S-NCP. The novel finding we report is that the early stage of 601-NCP dissociation involves a rapid asymmetric release of the DNA. After about half of the 601 DNA is released, proteins in a disrupted histone core remain trapped in a 'J'-shaped DNA conformation for at least 200 ms. There is evidence to support the biological relevance of the “semi-open” conformation described here by SAXS and previously by FRET [7]. This species may be the preferred substrate for histone chaperones and ATP-dependent chromatin remodeling complexes in transcriptional regulation as well as DNA replication and repair. The population of this intermediate may be regulated by post-translational modifications of the histones or incorporation of histone variants, like H2A.bbd that alter wrapping of DNA [39] or stabilize the NCP to histone exchange, like macroH2A [40]. More than 5 seconds pass before the histones detach from 601 DNA, perhaps a result of strong electrostatic interactions between DNA and the histones. In the 5S NCP, rapid asymmetric DNA

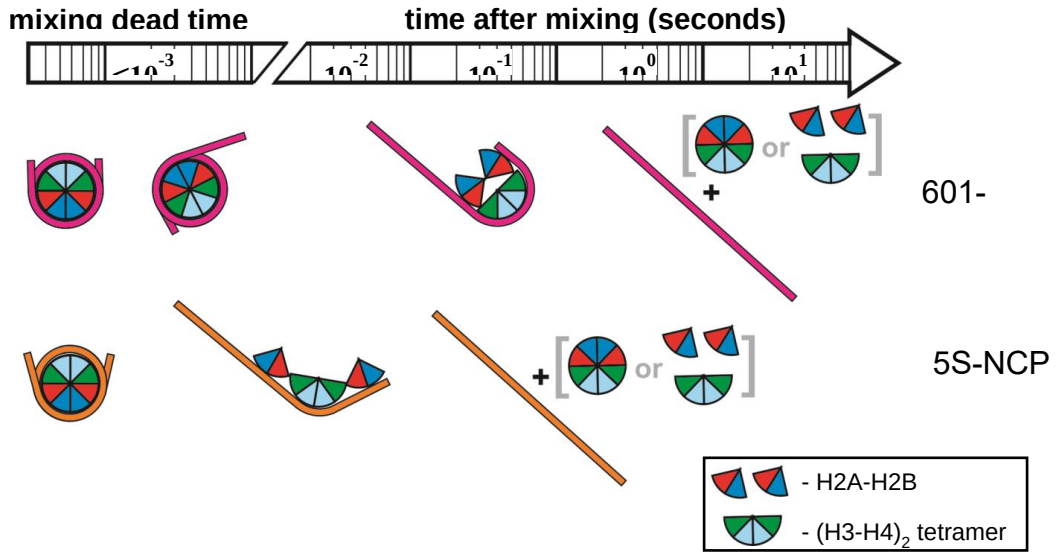


Figure 3.7 Timeline of salt-induced disassembly of 601-NCP and 5S-NCP**.** For 601-NCP, DNA opens rapidly from one end and reaches a metastable conformation within the first 20 ms. The histone octamer is disrupted by the asymmetric unwinding, but retains strong electrostatic interactions with 601 DNA and maintains a 'J'-shaped structure for ≈ 200 ms. This long-lived intermediate then dissociates at a rate of 0.74 ± 0.08 s^{-1}. 5S-NCPs exhibit much faster dissociation dynamics. After 20 ms, the DNA is mostly unwrapped and extended but still bound to the histone components. No stable intermediates are detected and 5S-NCPs disassemble within 1 s (two rates measured: 41.6 ± 13.9 s^{-1} and 1.13 ± 0.74 s^{-1}).

release occurs, followed almost immediately by protein dissociation. The absence of a long-lived intermediate precludes a more detailed analysis.

3.5 Conclusions

Our results establish a powerful platform for studying the global dynamics of nucleosomes and other nucleoprotein complexes that can be triggered by mixing. Technological advancements (i.e. brighter x-ray sources, faster high-viscosity mixers) will soon enable sub-millisecond studies, the timescale of large-scale dynamics for 5S-

NCP. The range of potential targets for this technique is broad and includes NCP variants as well as other protein-nucleic acid systems, including RNA-protein machines.

3.6 Acknowledgements

We thank Arthur Woll and CHESS staff for experimental assistance. We thank John Sonewald and DJ Black from BioLogic USA for technical assistance. We acknowledge support from NIH through EUREKA grant [R01-GM088645] and grant [R01-GM085062] to LP. LMG acknowledges support from NIH grant [GM073787]. Additional support was awarded to JMT from NIH grant [T32GM008267]. CHESS is supported by the NSF & NIH/NIGMS via NSF award [DMR-0936384].

REFERENCES

[1] Y. Chen, J.M. Tokuda, T. Topping, J.L. Sutton, S.P. Meisburger, S. A. Pabit, L.M. Gloss, L. Pollack, Revealing transient structures of nucleosomes as DNA unwinds, Nucleic Acids Res. 42 (2014) 8767–8776.

[2] K. Luger, S.E. Phillips, Rise of the molecular machines, Curr. Opin. Struct. Biol. 20 (2010) 70–72.

[3] T.D. Yager, C.T. McMurray, K.E. van Holde, Salt-induced release of DNA from nucleosome core particles., Biochemistry. 28 (1989) 2271–2281.

[4] J.L. Killian, M. Li, M.Y. Sheinin, M.D. Wang, Recent advances in single molecule studies of nucleosomes., Curr. Opin. Struct. Biol. 22 (2012) 80–7.

[5] A. Gansen, A. Valeri, F. Hauger, S. Felekyan, S. Kalinin, K. Tóth, J. Langowski, C. A. M. Seidel, Nucleosome disassembly intermediates characterized by single-molecule FRET., Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 15308–15313.

[6] V. Böhm, A.R. Hieb, A.J. Andrews, A. Gansen, A. Rocker, K. Tóth, K. Luger, J. Langowski, Nucleosome accessibility governed by the dimer/tetramer interface, Nucleic Acids Res. 39 (2011) 3093–3102.

[7] D.A. Hoch, J.J. Stratton, L.M. Gloss, Protein-Protein Förster Resonance Energy Transfer Analysis of Nucleosome Core Particles Containing H2A and H2A.Z, J. Mol. Biol. 371 (2007) 971–988.

[8] W.J. a Koopmans, a. Brehm, C. Logie, T. Schmidt, J. Van Noort, Single-pair FRET microscopy reveals mononucleosome dynamics, J. Fluoresc. 17 (2007) 785–795.

[9] M. Tomschik, K. Van Holde, Nucleosome Dynamics as Studied by Single-pair Fluorescence Resonance Energy Transfer: A Reevaluation, J. Fluoresc. 19 (2009) 53–62.

[10] J. Zlatanova, T.C. Bishop, J.M. Victor, V. Jackson, K. van Holde, The Nucleosome Family: Dynamic and Growing, Structure. 17 (2009) 160–171.

[11] G. Li, M. Levitus, C. Bustamante, J. Widom, Rapid spontaneous accessibility of nucleosomal DNA., Nat. Struct. Mol. Biol. 12 (2005) 46–53.

[12] H.S. Tims, K. Gurunathan, M. Levitus, J. Widom, Dynamics of nucleosome invasion by DNA binding proteins, J. Mol. Biol. 411 (2011) 430–448.

[13] Y. Arimura, H. Tachiwana, T. Oda, M. Sato, H. Kurumizaka, Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome, Biochemistry. 51 (2012) 3302–3309.

[14] M.L. Kireeva, W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, V.M. Studitsky, Nucleosome Remodeling Induced by RNA Polymerase II, Mol. Cell. 9 (2002) 541–552.

[15] A. Miyagi, T. Ando, Y.L. Lyubchenko, Dynamics of nucleosomes assessed with time-lapse high-speed atomic force microscopy., Biochemistry. 50 (2011) 7901–7908.

[16] A.J. Andrews, K. Luger, A Coupled Equilibrium Approach to Study Nucleosome Thermodynamics, in: Methods Enzymol., 2011: pp. 265–285.

[17] A.J. Andrews, K. Luger, Nucleosome structure(s) and stability: variations on a theme., Annu. Rev. Biophys. 40 (2011) 99–117.

[18] Y.J. Park, P.N. Dyer, D.J. Tremethick, K. Luger, A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome, J. Biol. Chem. 279 (2004) 24274–24282.

[19] J. Mazurkiewicz, J.F. Kepert, K. Rippe, On the mechanism of nucleosome assembly by histone chaperone NAP1, J. Biol. Chem. 281 (2006) 16462–16472.

[20] P.T. Lowary, J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning., J. Mol. Biol. 276 (1998) 19–42.

[21] R.T. Simpson, F. Thoma, J.M. Brubaker, Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: A model system

for study of higher order structure, Cell. 42 (1985) 799–808.

[22] D.I. Svergun, M.H.J. Koch, Small-angle scattering studies of biological macromolecules in solution, Reports Prog. Phys. 66 (2003) 1735–1782.

[23] C. Yang, M.J. Van Der Woerd, U.M. Muthurajan, J.C. Hansen, K. Luger, Biophysical analysis and small-angle X-ray scattering-derived structures of MeCP2-nucleosome complexes, Nucleic Acids Res. 39 (2011) 4122–4135.

[24] S. Mangenot, A. Leforestier, P. Vachette, D. Durand, F. Livolant, Salt-induced conformation and interaction changes of nucleosome core particles., Biophys. J. 82 (2002) 345–56.

[25] A. Bertin, A. Leforestier, D. Durand, F. Livolant, Role of Histone Tails in the Conformation and Interactions of Nucleosome Core Particles, Biochemistry. 43 (2004) 4773–4780.

[26] A. Bertin, M. Renouard, J.S. Pedersen, F. Livolant, D. Durand, H3 and H4 histone tails play a central role in the interactions of recombinant NCPs., Biophys. J. 92 (2007) 2633–2645.

[27] A. Bertin, D. Durand, M. Renouard, F. Livolant, S. Mangenot, H2A and H2B tails are essential to properly reconstitute nucleosome core particles, Eur. Biophys. J. 36 (2007) 1083–1094.

[28] Y. Inoko, M. Yamamoto, F. Satoru, T. Ueki, X-Ray Scattering Study of the Shape of the DNA Region Core Particle with Synchrotron Radiation, J. Biochem. 111 (1992) 310–316.

[29] D.I. Svergun, M.H.J. Koch, Structural Model of the 50 S Subunit of Escherichia coli Ribosomes from Solution Scattering, J. Mol. Biol. (1994) 66–77.

[30] P.N. Dyer, R.S. Edayathumangalam, C.L. White, Y. Bao, S. Chakravarthy, U.M. Muthurajan, K. Luger, Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA, Methods Enzymol. 375 (2004) 23–44.

[31] A. Thaström, L.M. Bingham, J. Widom, Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning, J. Mol. Biol. 338 (2004) 695–709.

[32] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.

[33] R. Russell, I.S. Millett, M.W. Tate, L.W. Kwok, B. Nakatani, S.M. Gruner, S.G.J. Mochrie, V. Pande, S. Doniach, D. Herschlag, L. Pollack, Rapid compaction during RNA folding, Proc. Natl. Acad. Sci. 99 (2002) 4266–4271.

[34] T.J. Macke, D.A. Case, Modeling Unusual Nucleic Acid Structures, in: N.B. Leontis, J. Santalucia (Eds.), Mol. Model. Nucleic Acids, American Chemical Society, Washington, DC, 1998: pp. 379–393.

[35] D. Svergun, C. Barberato, M.H. Koch, CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 28 (1995) 768–773.

[36] S.C. Howell, K. Andresen, I. Jimenez-Useche, C. Yuan, X. Qiu, Elucidating internucleosome interactions and the roles of histone tails, Biophys. J. 105 (2013) 194–199.

[37] P. Bernadó, E. Mylonas, M. V Petoukhov, M. Blackledge, D.I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering., J. Am. Chem. Soc. 129 (2007) 5656–64.

[38] E.Y.D. Chua, D. Vasudevan, G.E. Davey, B. Wu, C.A. Davey, The mechanics behind DNA sequence-dependent properties of the nucleosome., Nucleic Acids Res. 40 (2012) 6338–52.

[39] Y. Bao, K. Konesky, Y.-J. Park, S. Rosu, P.N. Dyer, D. Rangasamy, D.J. Tremethick, P.J. Laybourn, K. Luger, Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA., EMBO J. 23 (2004) 3314–3324.

[40] S. Chakravarthy, S.K.Y. Gundimella, C. Caron, P. Perche, J.R. Pehrson, S. Khochbin, Structural Characterization of the Histone Variant macroH2A Structural Characterization of the Histone Variant macroH2A, Mol. Cell. Biol. 25 (2005) 7616–7624.

Chapter 4

Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core †

Abstract

The nucleosome core particle (NCP) is the basic structural unit for genome packaging in eukaryotic cells and consists of DNA wound around a core of eight histone proteins. DNA access is modulated through dynamic processes of NCP disassembly. Partly-disassembled structures, such as the hexasome (containing six histones) and the tetrasome (four histones), are important for transcription regulation in vivo. However, the pathways for their formation have been difficult to characterize. We combine time-resolved SAXS (TR-SAXS) and time-resolved FRET (TR-FRET) to correlate changes in the DNA conformations with composition of the histone core during salt-induced disassembly of canonical NCPs. We find that H2A-H2B histone dimers are released sequentially, with the first dimer being released after the DNA has formed an asymmetrically unwrapped, "teardrop"-shape DNA structure. This finding suggests that the octasome-to-hexasome transition is guided by the asymmetric unwrapping of the DNA. The link between DNA structure and histone composition suggests a

---

† **This chapter is reprinted with permission from** [1]**:** Y. Chen*, J.M. Tokuda*, T. Topping, S.P. Meisburger, S.A. Pabit, L.M. Gloss, L. Pollack, Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core, Proc. Natl. Acad. Sci. 114 (2017) 334–339. It has been modified to fit the format of this thesis. *These authors contributed equally

potential, new mechanism for the action of proteins that alter nucleosome configurations such as histone chaperones and chromatin remodeling complexes.

4.1 Significant Statement

Nucleosomes are fundamental protein-DNA structures through which eukaryotes package and organize DNA inside the nucleus. Nucleosomes are disassembled to gain access to the critical information stored in DNA. Here, we describe a new experimental approach that characterizes the kinetics of nucleosome disassembly and the synergy between DNA conformation and protein components. Using NaCl to disrupt electrostatic interactions, we identify kinetic pathways and transient intermediates that reveal how DNA unwrapping and protein dissociation are linked in this macromolecular complex. These dynamic structures may provide new insight into the regulation of DNA access during transcription, replication and repair.

4.2 Introduction

Genome access is highly regulated through the hierarchical organization of proteins and nucleic acids within the cell nucleus. The nucleosome core particle (NCP) is the first level of this hierarchy [2] and contains two dimers of H2A-H2B histones and an $(H3\text{-}H4)_2$ tetramer that is assembled as a dimer-of-dimers. Around this symmetric octamer core, ≈ 146 base pairs of dsDNA are wrapped in ≈ 1.7 super-helical turns [2,3]. The NCP structure physically impedes access to DNA, but is dynamically modulated by numerous mechanisms: post-translational modification (PTM) of histones, incorporation of histone variants, DNA sequence dependent effects and the actions of

extrinsic protein factors (e.g. histone chaperones, ATP-dependent remodeling complexes, and histone PTM binding proteins) [4,5].

Studies of the intrinsic properties and dynamics of NCPs are critical for understanding how nuclear machinery gains DNA access in vivo [4,6,7]. Insight into the nature of partially unfolded NCP structures has been gleaned from in vitro studies of NCP assembly and disassembly. Intermediate species with partially unwrapped DNA [8,6], disrupted histone-histone interfaces [9,10], and dissociation of one (hexasomes) or two (tetrasomes) H2A-H2B dimers have been reported [11–13]. Some of these NCP intermediates have been directly connected to chromatin function *in vivo*. For example, the hexasome is formed by the action of RNA Pol II [14] and the essential histone chaperone FACT [15].

In addition to equilibrium studies, the kinetics of nucleosome assembly and disassembly have been characterized by bulk and single-molecule methods, including Förster Resonance Energy Transfer (FRET) [8,9,16–18], atomic force microscopy (AFM) [10,19], force spectroscopy [20–22], and small angle x-ray scattering (SAXS) [11,23]. Many studies focused primarily on specific DNA-histone contacts and local conformational changes. Few, if any studies, employ complementary methods to directly compare, on similar kinetic time scales, the structural changes of the DNA and histone components of the NCP. A major gap in our understanding of NCP disassembly arises from our limited knowledge of the coordination between DNA conformation and histone core composition.

Since the NCP protein-DNA complex is stabilized predominantly by polycation-polyanion interactions, the in vitro equilibrium and kinetic properties can be

manipulated by ionic solvent conditions. NaCl has been widely used to study partially assembled, biologically relevant NCP species that are marginally populated under physiological conditions [6,17,24,25]. The use of recombinant histones and the Widom 601 DNA sequence (selected for its ability to form stable, well-positioned nucleosomes [26,27]) allows production of large amounts of homogeneous NCPs (601-NCP) for biophysical studies. Figure 4.1 shows the NaCl-induced disassembly pathway for 601-NCPs [8,9,24,25]. While the various species shown in Figure 4.1 have been detected at equilibrium, much less is known about the kinetics of NCP disassembly, including the relevant transition times and pathways between states, or the potential for coordination between DNA unwrapping and disruption of histone-histone interfaces.

Our recent time-resolved small angle x-ray scattering (TR-SAXS) study of salt-induced NCP dissociation revealed asymmetric DNA release from the histone octamer [23]. In kinetic jumps from ≈ 0 to 1.9 M NaCl, a transient intermediate was observed with the DNA in a "J"-shape conformation bound to a disrupted histone core. We applied contrast variation TR-SAXS to focus solely on the dynamic changes in DNA conformation [23]. Information about the histone proteins was restricted to the "endpoint states" of intact or completely dissociated octamer.

Here, we report the coupling of TR-SAXS studies of DNA conformational changes with time-resolved FRET (TR-FRET) studies of H2A-H2B dimer dissociation during salt-induced NCP disassembly. Two conditions are characterized here: complete NCP disassembly following rapid increase from low salt (≈ 0) to 1.9 M NaCl (as in [23]), and partial disassembly upon increase to 1.2 M NaCl. The latter condition allowed observation of DNA conformations that facilitate release of the H2A-H2B dimers. The

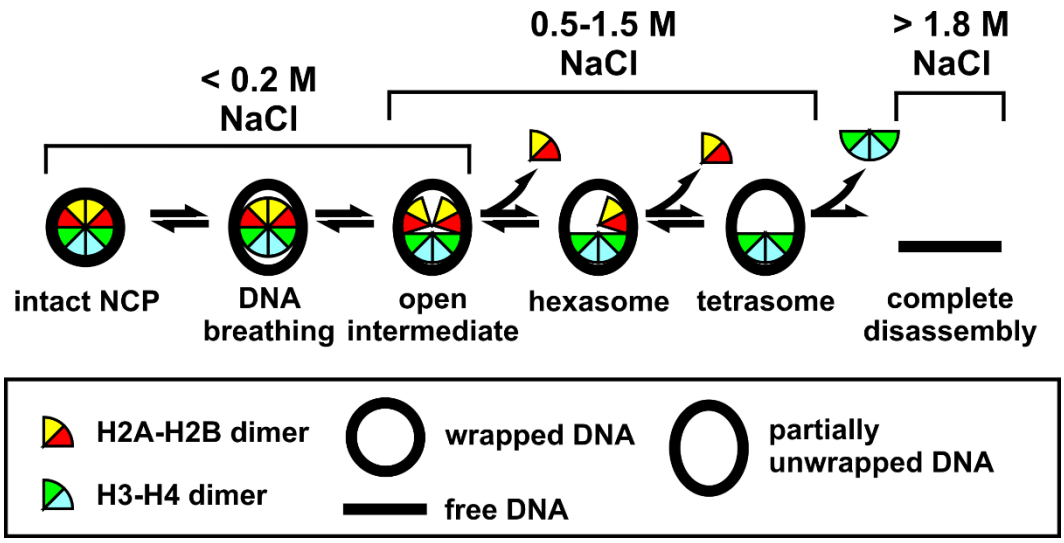


Figure 4.1 A schematic of NaCl-dependent disassembly for NCPs containing the 601 DNA [16], based on equilibrium studies ([NCP] ≥ 25 nM). At physiological ionic strength, NCP configurations reflect local dynamics (i.e. DNA breathing [7], and formation of an open intermediate [9]). Above 0.5 M NaCl, H2A-H2B dimers begin to dissociate, allowing the formation of hexasomes and tetrasomes [24]. Above 1.4 M NaCl, $(H3\text{-}H4)_2$ tetramers begin to dissociate, allowing for complete disassembly [25]. Although dimer dissociation is reversible, tetrasomes are the minimal configurations required to maintain a wrapped DNA structure.

combination of TR-SAXS and TR-FRET provides new insights into the conformational dynamics of open intermediate and hexasome formation (Figure 4.1), with important implications for the biological function of the NCP in regulating DNA accessibility.

4.3 Materials and Methods

4.3.1 NCP Production and Reconstitution

Previously described procedures were used for the expression and purification of recombinant *Xenopus laevis* histones [16,28,29], production of the 149 bp DNA derived from the Widom 601 sequence [16,26] and their reconstitution into nucleosome core

particles (NCP) [23]. Unless noted otherwise, experiments were conducted with the following buffer: 20 mM Tris-Cl (pH 7.5), 0.1 mM EDTA, 0.1 mM DTT.

4.3.2 TR-SAXS Experiments

All TR-SAXS experiments were conducted using a Biologic SFM-400 stopped flow mixer at BioCAT Sector 18 at Advanced Photon Source (APS). The experimental procedures and SAXS image analysis are described in detail in Appendix B.

4.3.3 Ensemble Optimization Method (EOM)

Ensembles of DNA structures that best recapitulate the measured TR-SAXS profiles were selected using the program GAJOE v2.0 [30,31]. The DNA pool included 9182 structural models generated using PyMol (expanded from 32 models in [23]). Details on model generation are described in Appendix B. The q-range used for GAJOE fitting was 0.006-0.2 $Å^{-1}$. Rg histograms of selected DNA models were averaged from 100 independent repeats of the genetic algorithm. Parameter settings: number of harmonics = 50, maximum s-value = 0.25, number of points = 101, number of generations = 10000, number of ensembles = 200, ensemble size fixed = no, maximum/minimum number of curves per ensemble = 1, constant subtraction = no, number of times genetic algorithm repeated = 100.

The reliability and uniqueness of the solutions achieved by the genetic algorithm depend on two interdependent factors: (1) the size and diversity of the pool (which needs to contain an ensemble that fits the data well, e.g. $\chi^2 \leq 2$) and (2) the number of generations and iterations of the algorithm (to provide sufficient sampling and evolution

in order to find the best fitting ensemble). The solutions obtained for a given SAXS profile using the DNA pool and parameter settings described above consistently converged to give nearly identical ensembles (with 0-2 bp variations).

4.3.4 TR-FRET Experiments

A previous paper described the design of a FRET system to monitor interactions in the NCP specifically between the H2A-H2B dimers, with Cys-AEDANS acceptors, and the $(H3\text{-}H4)_2$ tetramer, with Trp donors [16]. See details in Appendix B.

4.4 Results

4.4.1 DNA unwrapping at 1.9 M NaCl visualized by TR-SAXS

TR-SAXS with contrast variation was used to monitor the DNA conformations during complete NCP disassembly upon the rapid shift from ≈ 0 to 1.9 M NaCl by stopped-flow mixing (for SAXS profiles, see Supplementary Figure B.1A). In standard SAXS measurements, both protein and DNA contribute to the scattering. Interpretation of these SAXS profiles is challenging and requires knowledge of how each component is distributed. Through contrast variation, scattering from the DNA is isolated by matching the electron density of the solvent to that of the protein. As illustrated in Figure 4.2, this condition is achieved through the addition of 50% sucrose to the buffer. This contrast matched condition allows for unambiguous analysis of the DNA conformation since only the DNA contributes to the SAXS signal. Sucrose is an effective contrast additive since it negligibly affects electrostatics [32] and NCP

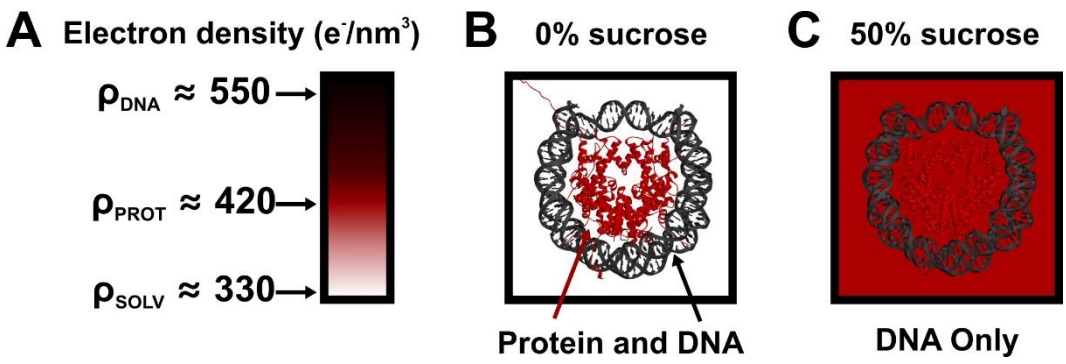


Figure 4.2 Contrast variation SAXS isolates structural information for the DNA component of NCPs. (**A**) A color scale bar with typical electron density values for solvent (water), protein, and DNA. (**B-C**) NCP structures (PDB 1AOI) shown in buffers with electron densities that vary depending on the presence of 0% (**B**) or 50% (**C**) sucrose. We used contrast variation SAXS to monitor DNA conformations during NCP disassembly induced with a salt-jump.

stability [23].

Since previous equilibrium and time-resolved SAXS studies of NCP disassembly revealed the presence of multiple structures, an ensemble optimization method (EOM) was applied to identify ensembles of DNA conformations that best recapitulate the SAXS profiles [23,30,31]. An overview of this experimental strategy is summarized in Figure 4.3. A pool of 9,182 nucleosomal DNA structures was generated with varying degrees of DNA unwrapping to create a large conformational space. For each time point, a genetic algorithm selected a subset of structures (or "ensemble") that yields a theoretical SAXS profile that best fits the SAXS data (for details, see Materials and Methods; for χ^2-values and example fit, see Supplementary Figure B.2A-B). The optimized ensemble consists of representative structures that closely approximate the conformations in solution.

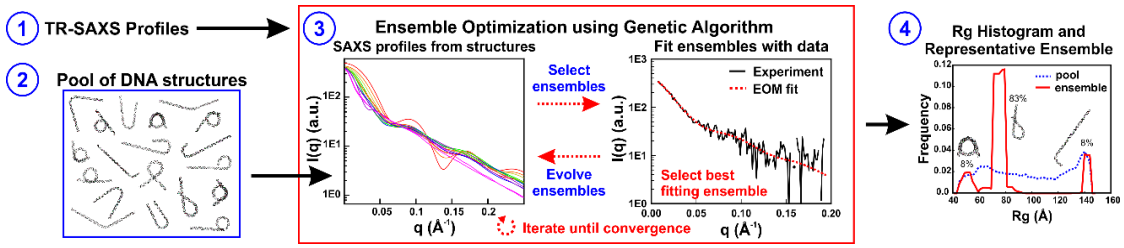


Figure 4.3 Overview of the ensemble optimization method (EOM) used for determining structures. Ensemble optimization (Step 3, red box) requires SAXS profiles (Step 1) and a pool of DNA structures (Step 2) that contains a large number of possible conformations. First, the theoretical SAXS profile for each structure in the pool is calculated using CRYSOL (Step 3, *left*). A genetic algorithm (GAJOE) randomly selects subsets of these structures, called ensembles, for comparison with the input SAXS data (Step 3, *right*). Structures from the best fitting ensembles are propagated into the next generation of ensembles along with some new structures, and this process is repeated (10,000 times) until convergence is achieved. The entire ensemble optimization process is repeated (100 times) to confirm reproducibility and the final ensembles that best represent the data are used to generate histograms of the radius of gyration, and to determine the most representative structures for the SAXS profiles (Step 4). The example fit and results shown is for the 300 ms time point of NCPs in 1.2 M NaCl under contrast-matched conditions (proteins "invisible" in 50% sucrose).

To visualize the ensembles, we calculated size distributions (radius of gyration, Rg) of the structures selected by EOM for each time point (Figure 4.4A; for initial DNA pool, see Supplementary Figure B.2C). Those with an Rg of ≈ 50 Å represent mostly–wrapped DNA. The diminishing amplitude of this peak with time corresponds to the disruption of the canonical NCP structure and the population of other conformations. Population of a state with an Rg near 140 Å (the Rg for a fully extended DNA) represents complete disassembly. Structures with intermediate values of Rg (between the peaks at 50 Å and 140 Å) represent intermediate states. The presence of multiple peaks at each time point suggests heterogeneous populations of NCP structures. The dominant conformations selected by EOM for the fully wrapped and extended states are shown in Figure 4.4B and the major conformations selected for the intermediate states

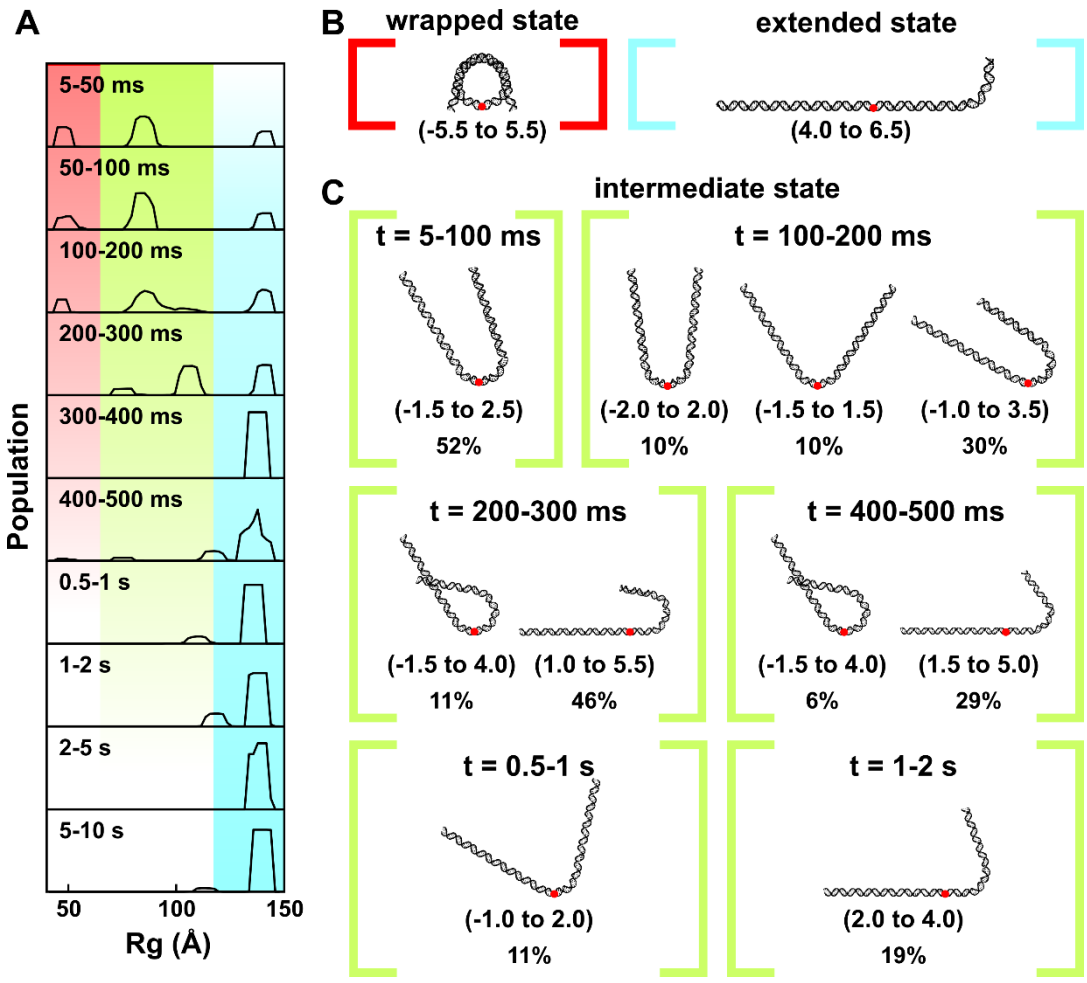


Figure 4.4 DNA structures selected by ensemble optimization method (EOM) analysis of TR-SAXS data for NCPs dissociated in 1.9 M NaCl and 50% sucrose**.** (**A**) Rg(t) histograms for DNA structural models selected by EOM. Regions highlighted in red, green and blue correspond with the fully wrapped, intermediate and extended states, respectively. (**B**) Models that best represent the measured SAXS profile for the initial wrapped state (red) and final extended state (blue). (**C**) Models that best represent the intermediate states as a function of time. Red dots indicate the dyad axis or superhelical location zero (SHL 0). Numbers in the parentheses reveal the range of SHLs (number of turns where the major groove faces the histone, every 10 bp) contained within the curved portions. Percentages show the weights of the species out of the total population at the indicated time point. Under high-salt conditions where complete dissociation of 601-NCPs is favored, multiple partially unwrapped intermediates are populated.

are shown in Figure 4.4C. Within 100 ms, approximately half of the NCPs contain nearly symmetrically unwrapped DNA in a "U-shape," a quarter of the DNAs are fully wrapped, and a quarter of the DNAs are fully unwrapped. Between 100-500 ms, the partially wrapped intermediate states become more asymmetric with most of the population exhibiting "J"-shaped structures. A small fraction form "teardrop" structures, where one end of the DNA remains wrapped around the histone core, while the other is extended. After 2 s, the nucleosomal DNA is predominantly unwrapped.

4.4.2 NCP dissociation at 1.2 M NaCl visualized by TR-SAXS

To better characterize transient, asymmetric DNA species, we applied the same approach to study NCP disassembly at a final NaCl concentration of 1.2 M (for SAXS profiles, see Supplementary Figure B.1B). This lower salt concentration limits disassembly to tetrasome species (tetramer dissociation occurs above 1.4 M NaCl, see Figure 4.1, [16,25]), and captures DNA conformations associated with release of the H2A-H2B dimers.

Compared to 1.9 M NaCl, NCP unwrapping at 1.2 M NaCl was significantly slower and was incomplete within our 10 s measurement window. The DNA ensembles selected by EOM (Supplementary Figure B.3A) are quantified in the Rg histograms shown in Figure 4.5A. A fully unwrapped state (Rg ≈ 140 Å) appears after 300 ms, but comprises only about 8% of the ensemble. In a majority of the NCPs, the DNA remains partially wrapped with Rgs that range from 50 Å to 90 Å. For the first 200 ms of the reaction, these partially wrapped species vary in size and shape; after 300 ms these structures converge to one with an Rg of 76 Å (circled in red). Figure 4.5B shows

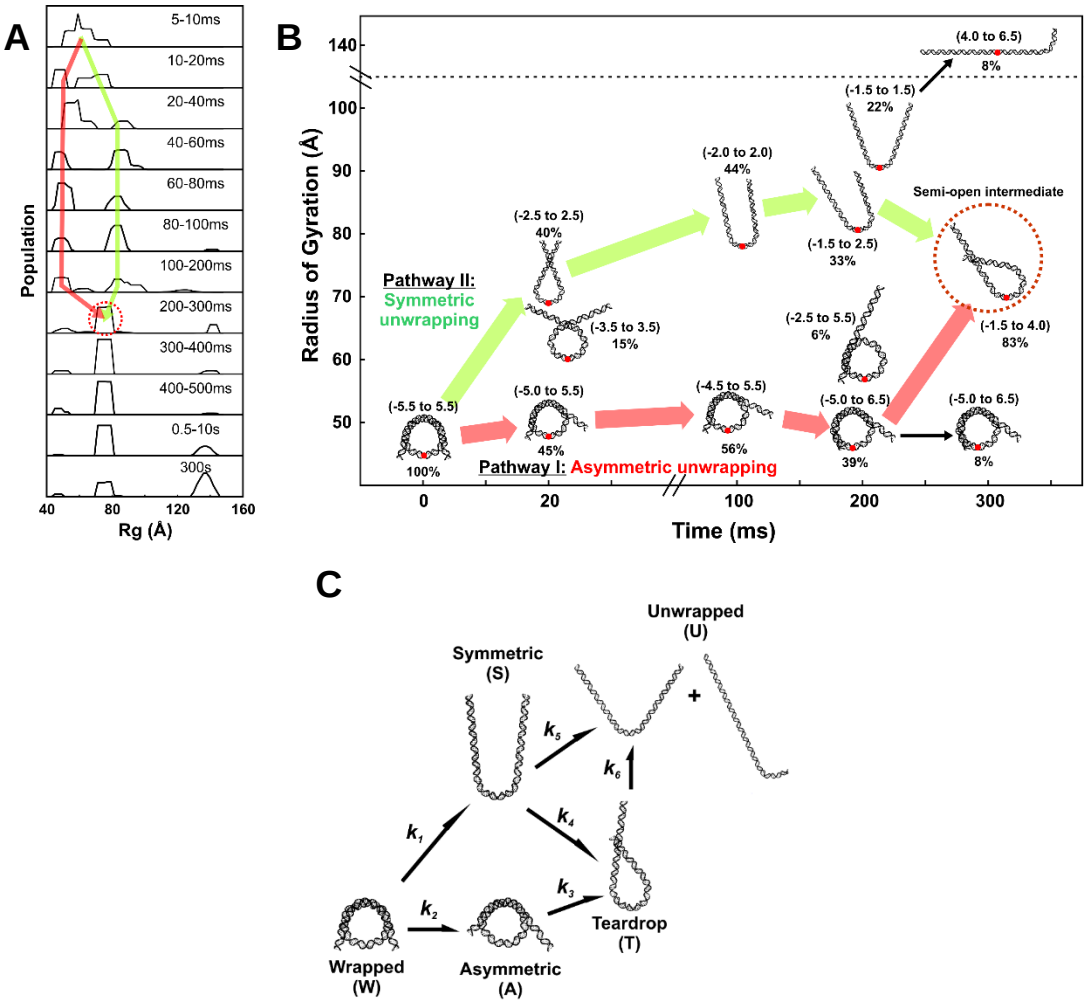


Figure 4.5 DNA structures selected by EOM for NCPs dissociated in 1.2 M NaCl. (**A**) Rg histograms from DNA models selected by EOM that best represent the SAXS data. Red and green arrows highlight two pathways through which DNA structures change before settling into a prominent peak after 300 ms (circled in red). (**B**) DNA models selected by EOM before (t = 0) and after mixing into 1.2 M NaCl (20 ms, 100 ms, 200 ms, 300 ms). Green and red arrows highlight two major pathways through which DNA unwraps to form the "teardrop" DNA structure. Black arrows show minor pathways. Red dots indicate the dyad axis (SHL 0). Numbers in the parentheses reveal the range of SHLs (number of turns) contained within the curved portions. Percentages shown are the weights of the species out of the total population at the indicated time point. Under moderate salt conditions that favor partial disassembly, the majority of structures unwrap symmetrically and asymmetrically before converging into the teardrop structure. (**C**) Kinetic scheme for complete disassembly with pathways inferred from prominent DNA structures selected by EOM (Supplementary Figure B.3).

representative structures for the major populations for the first 300 ms, together with proposed pathways for the time-dependent evolution of DNA unwrapping. After 300 ms, the asymmetric "teardrop"-shaped DNA is the predominant species. As highlighted in Figure 4.5A-B, DNA reaches this structure through two major pathways. In Pathway I, the "teardrop" forms directly as DNA unwraps asymmetrically. In Pathway II, DNA initially unwraps symmetrically, but one end re-wraps to form the asymmetric intermediate. The "teardrop" is relatively stable at 1.2 M NaCl and represents 80% of the population at 300 ms and 36% of the population after 5 minutes (assessed from manual mixing experiments). The presence of two converging pathways was further confirmed by running another ensemble analysis algorithm: the minimal ensemble search (See Appendix B) [33].

To generate the kinetic pathway for complete disassembly, the DNA structures selected by EOM were binned into general classes of DNA structures (Supplementary Figure B.3B) and incorporated into the kinetic scheme shown in Figure 4.5C.

4.4.3 Kinetics of NCP Core Opening and Sequential H2A-H2B Dimer Release

FRET measurements that monitor the dissociation of H2A-H2B histone dimers from the octamer [16] provide complementary information to augment the DNA structures shown in Figure 4.4B-C and 4.5B. This system exploits unique Trp donors on the $(H3\text{-}H4)_2$ tetramer and acceptors on the H2A-H2B dimers, incorporated through modification of single Cys residues with IAEDANS (abbreviated CA). Since the NCP contains two copies of each core histone protein, each NCP contains two donors (D and

D') and two acceptors (A and A'). The sites for the FRET pairs (H3-78W donor to H2B-109CA acceptor) were chosen so that each FRET pair (i.e. the D-A and D'-A') contributes ≈ 50% to the overall FRET signal (Figure 4.6A), with minimal contribution from the other possible FRET combinations (D-A' and D'-A). Details of the FRET pairs and their influence on nucleosome stability have been thoroughly investigated in Ref. [16]. Previous FRET based studies have mostly focused on DNA unwrapping through the incorporation of FRET pairs between different positions on the DNA [17,18] or between the DNA and histones [8,9]. The advantage of this protein-protein FRET system is that it allows for unambiguous detection of H2A-H2B dimer release.

As shown in Figure 4.1, equilibrium studies utilizing FRET have identified at least three key intermediates that are populated at increasing NaCl concentrations: an open intermediate, the hexasome and the tetrasome. To assess the relevant time scales for histone dissociation, NCPs were rapidly shifted from ≈ 0 M into solutions with final NaCl concentrations that ranged from 0.7 to 2 M NaCl. In this survey of [NaCl]-dependence, the data were fit to a sum of exponentials and two major kinetic phases were observed for the loss of FRET (for details, see Appendix B and Supplementary Figure B.4). These phases were assigned to the formation of the hexasome and tetrasome (supported by native gel electrophoresis of NCPs incubated at 1 M NaCl for varying times (Supplementary Figure B.5)). There was evidence of a faster, minor kinetic phase (10-20% amplitude), on the 100 ms to 3 s time scale. However, this phase could not be quantitatively analyzed by the experimental approaches employed in this survey of NaCl conditions.

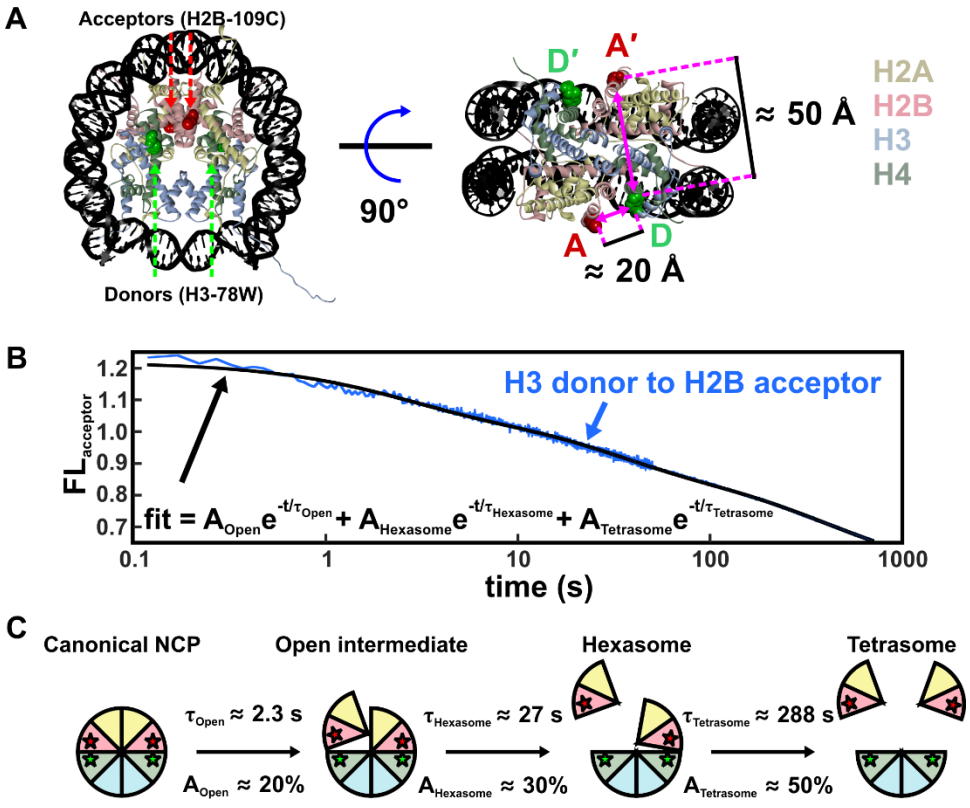


Figure 4.6 NCP FRET pairs and the histone configurations observed**.** (**A**) FRET pairs with H3-78W donor (green) and H2B-109CysAEDANS acceptor (red). For this construct (H3-H2B NCP), the donor and acceptor on the same face of the NCP (D-A) are close to the Förster radius for this FRET pair (≈ 20 Å), but the distance from the donor to the acceptor on the other NCP face (D-A') is significantly longer (≈ 50 Å) and should contribute less than 1% to the observed FRET signal. The Cβ positions in the 1AOI.pdb structure of the NCP were used to estimate distances between the FRET pairs. (**B**) Acceptor fluorescence time-course measured for 250 nM NCP in 1.2 M NaCl (blue). The solid black line represents a sum of three first-order exponentials used to determine the relative amplitudes and relaxation times. In order to obtain robust values, global fits were used on datasets collected as a function of NCP concentration (10-250 nM NCP). (**C**) Histone configurations observed with TR-FRET. Relaxation times (τ) and amplitudes (A) of FRET loss measured at 1.2 M NaCl are reported for each transition.

The relaxation times for the two major phases decrease monotonically, in a parallel pattern, with increasing NaCl concentrations (Supplementary Figure B.4). First, a monotonic decrease demonstrates that the kinetic pathways for dimer dissociation is relatively smooth across NaCl concentrations, which favor partial disassembly to the

tetrasome below 1.5 M and complete disassembly above 1.8 M. Thus, SAXS and FRET studies at two NaCl concentrations (1.2 M and 1.9 M) should provide a consistent kinetic model for NCP dissociation, with the caveat that intermediates are likely to be more stably populated at the lower NaCl concentration. Second, the parallel NaCl-dependence of these two major kinetic phases suggests that their transition states involve disruption of similar macromolecular interactions. Thus, these kinetic phases likely reflect similar reactions in a sequential mechanism (e.g. dissociation of the first and then second H2A-H2B dimer). This conclusion is supported by analysis of the relative FRET amplitudes from multiple FRET pairs described below.

To better characterize the nature of the kinetic phases, especially the fastest, low-amplitude kinetic phase, larger datasets were collected as a function of final NCP concentration (25 to 250 nM), at 1.2 M and 1.9 M NaCl, using a combination of manual and stopped-flow mixing to monitor reactions from 10 ms to 1000 s. Datasets at different NCP concentrations were globally fit to three kinetic phases. These results are shown in Figure 4.6B and are summarized in Table 4.1 (for details, see Appendix B and Supplementary Figure B.6A-B). The relative amplitudes for the three kinetic phases provide insight into the conformational changes associated with each phase as shown in Figure 4.6C. The relatively small amplitude of the fastest phase (loss of ≈ 20% of the FRET signal at both 1.2 M and 1.9 M) is consistent with a conformational change that opens the dimer-tetramer interfaces, rather than full dissociation of an H2A-H2B dimer. The larger amplitudes for the slower phases are consistent with dimer dissociation. Surprisingly, the relative amplitudes for the slower, dissociation phases are unequal (≈ 30% and ≈ 50%). This amplitude pattern is consistent with a sequential formation of an

Table 4.1. Comparison of relaxation times and relative amplitudes from FRET and SAXS.‡

Relaxation times (s) / Relative amplitude (%)			
1.2 M NaCl	H3-78W to H2B-109CA	H4-60W to H2A-108CA	SAXS
τ_{Open}/ A_{Open}	2.3 (0.5) / 20 (4)	2.6 (0.6) / 23 (3)	2.7
$\tau_{Hexasome}$ / $A_{Hexasome}$	27 (3) / 33 (4)	29 (3) / 47 (3)	14
$\tau_{Tetrasome}$ / $A_{Tetrasome}$	288 (30) / 47 (6)	188 (38) / 30 (5)	--
1.9 M NaCl			
τ_{Open} / A_{Open}	0.22 (0.07) / 23 (7)	0.06 (0.03) / 10 (8)	0.21
$\tau_{Hexasome}$ / $A_{Hexasome}$	1.6 (0.3) / 26 (5)	1.5 (0.2) / 58 (14)	1.2
$\tau_{Tetrasome}$ / $A_{Tetrasome}$	8.1 (0.8) / 50 (7)	3.7 (0.7) / 32 (16)	--

‡The errors associated with the kinetic parameters are indicated in the parentheses.

asymmetric open intermediate, in which only one dimer-tetramer interface is disrupted, followed by dissociation of this H2A-H2B dimer to form the hexasome and then dissociation of the second dimer to form the tetrasome (Figure 4.6C). A detailed comparison of this model with that expected for a symmetric formation of the open intermediate is shown in Supplementary Figure B.7A.

In order to verify the kinetic model presented in Figure 4.6C, kinetic parameters were measured using a second FRET pair (H4-60W donor to H2A-108CA acceptor, Supplementary Figure B.7B). Despite measuring different interactions, the relaxation times from the second FRET pair (H4-H2A NCP) agree well, and the inequality of the amplitudes for the hexasome and tetrasome phases is consistent with the two dimers being released through different pathways (Table 4.1, Supplementary Figure B.7, Appendix B). Table 4.1 also provides kinetic parameters determined from the SAXS data collected in the absence of sucrose (for details, see Supplementary Figure B.8). The reasonable agreement of relaxation times determined by FRET and SAXS is highlighted in the overlay shown in Supplementary Figure B.6C.

4.5 Discussion

Although nucleosome disassembly is crucial for DNA access, the dynamics of this process is largely unexplored. This study combines knowledge of the DNA conformations monitored by SAXS with insight into histone configurations reported by FRET to provide new details on dynamics and coordination between DNA and histone proteins during NCP disassembly.

4.5.1 Kinetic models for DNA unwrapping

Time-resolved SAXS revealed multiple pathways through which nucleosomal DNA unwraps during salt-induced disassembly. Although small populations of NCPs (< 25% at 1.9 M and < 10 % at 1.2 M) dissociate at a rate that exceeds our limit of detection, in the majority of cases, NCPs progressively unwrap from the ends with rates that increase as a function of [NaCl]. Following a jump to high salt (1.9 M NaCl), complete disassembly is achieved within 10 s. The jump to a lower final salt concentration (1.2 M NaCl) reveals significantly slower kinetics, with a majority of the DNAs remaining partially wrapped. Under both conditions, the DNA unwraps to form asymmetric conformations ("J" and "teardrop" shapes). These observations are consistent with previous studies of nucleosomes containing the non-palindromic 601 DNA sequence, where asymmetric nucleosome stability is observed with the 5' end showing a greater binding affinity than the 3' end [27]. Asymmetric unwrapping may be a generalizable feature of DNA sequences with asymmetric affinities, as observed for both the 601 and 5S DNA sequences [23].

The slower kinetic responses at 1.2 M NaCl reveal two pathways to form the

"teardrop" DNA. In addition to asymmetric unwrapping (Pathway I), the detection of symmetric unwrapping and re-wrapping of one end to form the same conformation (Pathway II) highlights the potential for coordination between the DNA ends. These results are consistent with that reported by Ngo and co-workers using force-FRET spectroscopy and Monte Carlo simulations [22]. Under low tension (< 5 pN), they report that both DNA ends unwrap and rewrap synchronously, but further opening of one end stabilizes the rewrapping of the other end in a manner that is directed by the local flexibility of the DNA. They suggest that the unwrapping of one end may help stabilize the wrapping of the other through an overall reduction in electrostatic repulsion. Interestingly, the rewrapping observed in our work is observed under conditions where electrostatic interactions should be effectively screened (1.2 M NaCl). One possible explanation is that the rewrapping is facilitated by the histone tails.

For some of the asymmetric models determined in this study, the curved portion of the DNA that remains in contact with the histone core appears shifted away from the dyad and closer to the entry/exit sites compared to canonical structures. This conformation may depend on a sliding of the histone-DNA contacts [34]. One intriguing possibility is that the partial DNA unwrapping may help facilitate nucleosome sliding.

4.5.2 Structures and Pathways of Hexasome Formation Suggest DNA Directed NCP Disassembly

The integrated results from SAXS and FRET at 1.2 M NaCl provide new insight into the transient species populated by 601-NCPs. The time-resolved DNA populations

(classified in Supplementary Figure B.3) were globally fitted to the kinetic scheme shown in Figure 4.5C to obtain relaxation times (Supplementary Figure B.9). These globally fitted populations are shown as a Figure 4.7A, along with the expected populations of the histone configurations based on the H3-H2B NCP FRET data in Table 4.1. The complete kinetic scheme of NCP disassembly at 1.2 M is presented in Figure 4.7B (for details, see Appendix B).

The DNA rapidly unwraps from the histone octamer to form the teardrop DNA, which is the dominant conformation on the 0.2 to 1 s timescale. This teardrop conformation forms appreciably faster than the 2 to 30 s required for the asymmetric opening that disrupts an interface between the $(H3\text{-}H4)_2$ tetramer and one of the H2A-H2B dimers and subsequent dissociation of the first dimer to form the hexasome. Such a state, containing partially unwrapped DNA, but a full complement of histone proteins, is completely consistent with results of Li et al [7], suggesting that contacts between the DNA and H2A-H2B dimers are disrupted by conformational dynamics observed under physiological conditions, yet other contacts prevent immediate release of one of the heterodimers. Thus, the teardrop DNA precedes changes in the histone octamer conformation, suggesting that the unwrapped DNA end acts like a gate to expose the proximal H2A-H2B dimer for release. The DNA further unwraps and releases the remaining dimer to form the tetrasome. This picture is in full agreement with equilibrium studies reported by Böhm et al [9].

This work suggests an intriguing new mechanism for NCP remodeling in which DNA conformation facilitates the reconfiguration of the histone core. Although the asymmetric nature of the DNA unwrapping and subsequent dimer dissociation observed

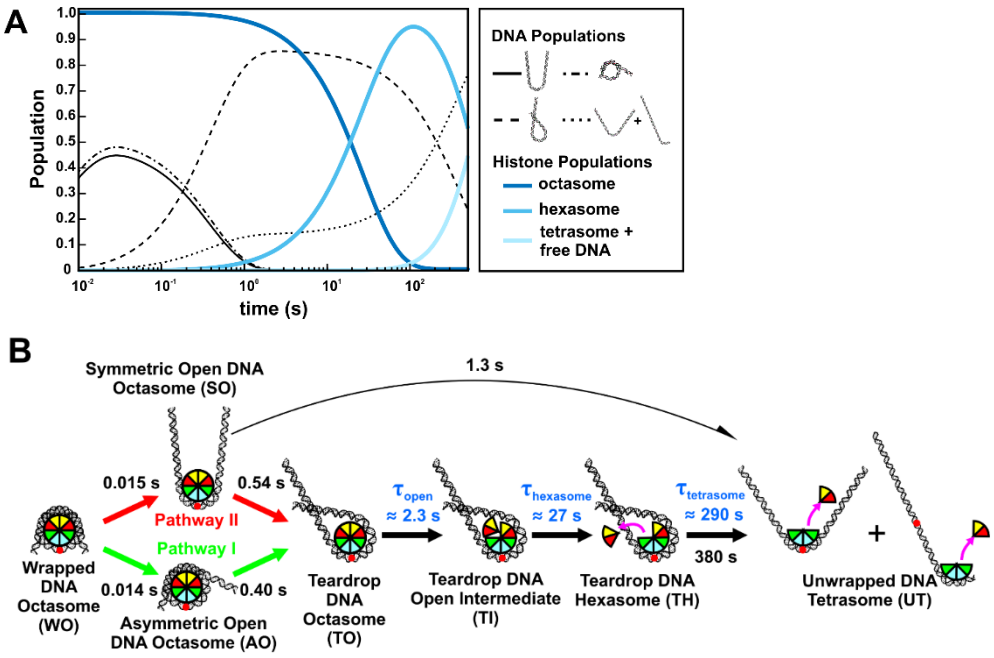


Figure 4.7 TR-FRET and TR-SAXS analyses reveal hexasome formation at 1.2 M NaCl. (**A**) Predicted populations of DNA conformational states (black lines) and histone configuration states (blue lines) based on the kinetic rates determined for NCPs at 1.2 M NaCl from the kinetic analysis of EOM models (Figure 4.6A, see details in Supplementary Figure B. 9) and TR-FRET measurements (Table 4.1), respectively. (**B**) NCP disassembly pathway determined from TR-SAXS with histone configurations informed by TR-FRET. Black numbers reflect the SAXS relaxation times (inverse of rates in Supplementary Figure B.9C). Blue numbers reflect the FRET relaxation times (Table 4.1). The curved black arrow represents a minor pathway. For simplicity, histone orientations were centered on the dyad when possible.

here was directed by the asymmetry of the tightly positioning 601 sequence, this mechanism may be exploited by gene regulatory proteins as a general strategy to exchange [35] or modify [36] one H2A-H2B dimer while simultaneously protecting the other. The combined SAXS and FRET approach used in this work is readily adaptable to test the hypothesis that, in addition to direct interaction with the histone core, key partner proteins, such as chromatin remodelers or histone chaperones, affect the composition of the histone core by interacting with and altering nucleosomal DNA conformation.

4.6 Conclusions

This work suggests an intriguing new mechanism for NCP remodeling in which DNA conformation facilitates the reconfiguration of the histone core. Although the asymmetric nature of the DNA unwrapping and subsequent dimer dissociation observed here was directed by the asymmetry of the tightly positioning 601 sequence, this mechanism may be exploited by gene regulatory proteins as a general strategy to exchange [35] or modify [36] one H2A-H2B dimer while simultaneously protecting the other. The combined SAXS and FRET approach used in this work is readily adaptable to test the hypothesis that, in addition to direct interaction with the histone core, key partner proteins, such as chromatin remodelers or histone chaperones, affect the composition of the histone core by interacting with and altering nucleosomal DNA conformation.

4.7 Acknowledgements

We thank Srinivas Chakravarthy, Weifeng Shang and Richard Heurich for experimental and technical assistance at BioCAT (Sector 18ID) Advanced Photon Source.

SAXS research was supported by National Institute of Health (NIH) grants EUREKA R01-GM088645 and R01-GM085062 to L.P. FRET research was supported by NIH GM073787 to L.M.G.

SAXS data were acquired at the BioCAT beamline of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the

DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. BioCAT is supported by grant 9 P41 GM103622 from the National Institute of General Medical Sciences (NIGMS) of the NIH. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the NIGMS or the NIH.

REFERENCES

[1] Y. Chen, J.M. Tokuda, T. Topping, S.P. Meisburger, S.A. Pabit, L.M. Gloss, L. Pollack, Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core, Proc. Natl. Acad. Sci. 114 (2017) 334–339.

[2] A.J. Andrews, K. Luger, Nucleosome structure(s) and stability: variations on a theme., Annu. Rev. Biophys. 40 (2011) 99–117.

[3] K. Luger, A.W. Mäder, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution., Nature. 389 (1997) 251–260.

[4] O. Bell, V.K. Tiwari, N.H. Thomä, D. Schübeler, Determinants and dynamics of genome accessibility., Nat. Rev. Genet. 12 (2011) 554–564.

[5] K. Luger, Dynamic nucleosomes., Chromosom. Res. 14 (2006) 5–16.

[6] H.S. Tims, K. Gurunathan, M. Levitus, J. Widom, Dynamics of nucleosome invasion by DNA binding proteins, J. Mol. Biol. 411 (2011) 430–448.

[7] G. Li, J. Widom, Nucleosomes facilitate their own invasion., Nat. Struct. Mol. Biol. 11 (2004) 763–769.

[8] G. Li, M. Levitus, C. Bustamante, J. Widom, Rapid spontaneous accessibility of nucleosomal DNA., Nat. Struct. Mol. Biol. 12 (2005) 46–53.

[9] V. Böhm, A.R. Hieb, A.J. Andrews, A. Gansen, A. Rocker, K. Tóth, K. Luger, J. Langowski, Nucleosome accessibility governed by the dimer/tetramer interface, Nucleic Acids Res. 39 (2011) 3093–3102.

[10] A. Miyagi, T. Ando, Y.L. Lyubchenko, Dynamics of nucleosomes assessed with time-lapse high-speed atomic force microscopy., Biochemistry. 50 (2011) 7901–7908.

[11] Y. Arimura, H. Tachiwana, T. Oda, M. Sato, H. Kurumizaka, Structural analysis of the hexasome, lacking one histone H2A/H2B dimer from the conventional nucleosome, Biochemistry. 51 (2012) 3302–3309.

[12] J. Mazurkiewicz, J.F. Kepert, K. Rippe, On the mechanism of nucleosome

assembly by histone chaperone NAP1, J. Biol. Chem. 281 (2006) 16462–16472.

[13] J. Zlatanova, T.C. Bishop, J.M. Victor, V. Jackson, K. van Holde, The Nucleosome Family: Dynamic and Growing, Structure. 17 (2009) 160–171.

[14] M.L. Kireeva, W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, V.M. Studitsky, Nucleosome Remodeling Induced by RNA Polymerase II, Mol. Cell. 9 (2002) 541–552.

[15] R. Belotserkovskaya, FACT Facilitates Transcription-Dependent Nucleosome Alteration, Science. 301 (2003) 1090–1093.

[16] D.A. Hoch, J.J. Stratton, L.M. Gloss, Protein-Protein Förster Resonance Energy Transfer Analysis of Nucleosome Core Particles Containing H2A and H2A.Z, J. Mol. Biol. 371 (2007) 971–988.

[17] N.P. Hazan, T.E. Tomov, R. Tsukanov, M. Liber, Y. Berger, R. Masoud, K. Toth, J. Langowski, E. Nir, Nucleosome Core Particle Disassembly and Assembly Kinetics Studied Using Single-Molecule Fluorescence, Biophys. J. 109 (2015) 1676–1685.

[18] Y.J. Park, P.N. Dyer, D.J. Tremethick, K. Luger, A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome, J. Biol. Chem. 279 (2004) 24274–24282.

[19] A.J. Katan, R. Vlijm, A. Lusser, C. Dekker, Dynamics of Nucleosomal Structures Measured by High-Speed Atomic Force Microscopy., Small. 11 (2014) 1–9.

[20] B.D. Brower-Toland, C.L. Smith, R.C. Yeh, J.T. Lis, C.L. Peterson, M.D. Wang, Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA., Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 1960–1965.

[21] M.A. Hall, A. Shundrovsky, L. Bai, R.M. Fulbright, J.T. Lis, M.D. Wang, High-resolution dynamic mapping of histone-DNA interactions in a nucleosome., Nat. Struct. Mol. Biol. 16 (2009) 124–129.

[22] T.T.M. Ngo, Q. Zhang, R. Zhou, J.G. Yodh, T. Ha, Asymmetric Unwrapping of Nucleosomes under Tension Directed by DNA Local Flexibility, Cell. 160 (2015) 1135–1144.

[23] Y. Chen, J.M. Tokuda, T. Topping, J.L. Sutton, S.P. Meisburger, S. a. Pabit, L.M. Gloss, L. Pollack, Revealing transient structures of nucleosomes as DNA unwinds, Nucleic Acids Res. 42 (2014) 8767–8776.

[24] I. Oohara, A. Wada, Spectroscopic Studies on Histone-DNA I. The Interaction interactions of Histone (H2A, H2B) Dimer with DNA:DNA Sequence Dependence, J. Mol. Biol. 196 (1987) 389–397.

[25] I. Oohara, A. Wada, Spectroscopic studies on histone-DNA interactions. II. Three transitions in nucleosomes resolved by salt-titration., J. Mol. Biol. 196 (1987) 399–411.

[26] P.T. Lowary, J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning., J. Mol. Biol. 276 (1998) 19–42.

[27] E.Y.D. Chua, D. Vasudevan, G.E. Davey, B. Wu, C.A. Davey, The mechanics behind DNA sequence-dependent properties of the nucleosome., Nucleic Acids Res. 40 (2012) 6338–52.

[28] D.D. Banks, L.M. Gloss, Equilibrium Folding of the Core Histones: the H3-H4 Tetramer Is Less Stable than the H2A-H2B Dimer, Biochemistry. 42 (2003) 6827–6839.

[29] L.M. Gloss, B.J. Placek, The effects of salts on the stability of the H2A-H2B histone dimer, Biochemistry. 41 (2002) 14951–14959.

[30] P. Bernadó, E. Mylonas, M. V Petoukhov, M. Blackledge, D.I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering., J. Am. Chem. Soc. 129 (2007) 5656–64.

[31] G. Tria, H.D.T. Mertens, M. Kachala, D.I. Svergun, Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering, IUCrJ. 2 (2015) 207–217.

[32] J.M. Blose, S. A. Pabit, S.P. Meisburger, L. Li, C.D. Jones, L. Pollack, Effects

of a protecting osmolyte on the ion atmosphere surrounding DNA duplexes, Biochemistry. 50 (2011) 8540–8547.

[33] M. Pelikan, G.L. Hura, M. Hammel, Structure and flexibility within proteins as identified through small angle X-ray scattering, Gen. Physiol. Biophys. 28 (2009) 174–189.

[34] G.D. Bowman, Mechanisms of ATP-dependent nucleosome sliding, Curr. Opin. Struct. Biol. 20 (2010) 73–81.

[35] P.J. Gonzalez, E. Palacian, Interaction of RNA Polymerase II with Structurally Altered Nucleosomal Particles, J. Biol. Chem. 264 (1989) 18457–18462.

[36] K. Yen, V. Vinayachandran, B.F. Pugh, XSWR-C and INO80 chromatin remodelers recognize nucleosome-free regions near +1 nucleosomes, Cell. 154 (2013) 1246–1256.

Chapter 5

Nucleotide-dependent unwrapping of the nucleosome by the Chd1 chromatin remodeler †

Abstract

Chromatin remodelers are ATP-dependent motors that reorganize DNA packaging by disrupting canonical histone-DNA contacts within the nucleosome. Here we show that the Chd1 chromatin remodeler stimulates DNA unwrapping from the edge of the nucleosome in a nucleotide-dependent fashion. As shown by FRET, DNA unwrapping by Chd1 readily occurred in the presence of ATP analogs ADP-BeF_3 and AMP-PNP, yet was also detectable in the absence of nucleotides. Unexpectedly, the DNA-binding domain of Chd1, which increases affinity and binds at the nucleosome edge, was not required for unwrapping to occur, suggesting that the ATPase domain disturbs histone-DNA contacts on the outer turn of DNA exiting the nucleosome. Using small angle x-ray scattering (SAXS) with contrast variation, we find that nucleosome unwrapping by Chd1 is highly asymmetric in solution, with different DNA conformations depending on the nucleotide-bound state. Ensemble modeling of SAXS data suggests that with ADP-BeF_3, Chd1 primarily promotes out-of-plane unwrapping from the histone core,

---

† **Manuscript in preparation.** Joshua M. Tokuda*, Ren Ren*, Robert F. Levendosky, Rebecca J. Tay, Ming Yan, Lois Pollack and Gregory D. Bowman
*Equal contribution.
Contributions: JMT - SAXS and analysis. RR and RFL – FRET, SFM, and analysis

whereas AMP-PNP complexes favor in-plane unwrapping. Taken together, our findings reveal previously unappreciated histone-DNA disruptions coupled to the nucleotide hydrolysis cycle of Chd1, which likely increase dynamics of nucleosome unwrapping during chromatin remodeling.

5.1 Introduction

Nucleosomes are the fundamental genomic packaging unit of eukaryotes, with approximately 146 bp of DNA tightly wrapped around a histone octamer [1]. In addition to compacting the genome into the small space of the nucleus, nucleosomes provide a means for storing epigenetic information while also regulating access to DNA. Chromatin remodelers are an essential class of enzymes that regulate DNA accessibility by disrupting the canonical histone-DNA interactions that occur in the nucleosome. Typically, remodelers alter the availability of DNA by assembling nucleosomes from histones and naked DNA, moving existing nucleosomes along DNA, or removing histones from DNA entirely [2]. Different families of remodelers appear to be specialized to achieve distinct biochemical outcomes, which correlate with their roles in vivo. Here, we focus on the Chd1 remodeler, which is involved in gene transcription and is required for stem cell pluripotency and suppression of cryptic transcription [3,4].

Chd1 has been found to act both in gene bodies and promoters, and directly interacts with a number of key factors involved in transcription. Chd1 influences nucleosome spacing in coding regions and directly interacts with several elongation factors such as Spt4-Spt5, the FACT complex, and the PAF complex [5–9]. In the absence of Chd1, chromatin appears to be perturbed by passage of RNA polymerase II (Pol II), which in

turn leads to cryptic transcription [4]. Thus, an important role of Chd1 is to help maintain the chromatin barrier as Pol II travels through the coding region.

Two primary biochemical activities have previously been demonstrated for Chd1: an ability to assemble nucleosomes and to reposition nucleosomes along DNA. Nucleosomes often fail to properly form when histones are rapidly deposited on naked DNA, and instead produce prenucleosomes, which lack the canonical superhelical wrapping of DNA around the histone core [10]. Chd1 catalyzes nucleosome assembly when histones are deposited on DNA by chaperones [11], which coincides with the ATP- and remodeler-dependent maturation of nucleosomes from prenucleosomes [10]. How nucleosome assembly is achieved is not yet understood, but has been shown to be a characteristic of ISWI-type remodelers as well.

The ability to reposition nucleosomes, also referred to as nucleosome sliding, is common to several remodeling families. For Chd1 and other remodelers such as SWI/SNF and ISWI, sliding results from local DNA translocation by the ATPase motor at an internal site on the nucleosome called superhelix location 2 (SHL2) [12–15]. Due to the inherent symmetry in the histone octamer and thus the nucleosome, there are two such SHL2 sites where ATPase motors can act, and two Chd1 ATPases can simultaneously bind to these two sites [16]. In addition to the ATPase motor, Chd1 possesses an N-terminal pair of chromodomains and a C-terminal DNA-binding domain (DBD) [17]. The chromodomains contact nucleosomal DNA at SHL1, adjacent to the ATPase motor, whereas the DBD binds the outer wrap of DNA at the edge of the nucleosome [16]. Due to the superhelical coiling of DNA around the histone core, the DBD is close to and can physically communicate with the chromo-ATPase on the inner

gyre of the nucleosome. Interestingly, both negative stain and cryoEM suggested that this organization of Chd1 domains correlates with partial unwrapping of the outer DNA gyre of the nucleosome [16,18].

Here we use fluorescence resonance energy transfer (FRET) and small angle x-ray scattering (SAXS) to study how Chd1 alters the organization of nucleosomal DNA. Unwrapping occurs with both AMP-PNP and ADP-BeF_3-bound states of the remodeler, and strongly points to the ATPase motor being the key element catalyzing this disruption of histone-DNA contacts. As shown by stopped-flow binding experiments, the association of Chd1 with the nucleosome is strongly nucleotide dependent, suggesting that these complex behaviors are due to distinct conformational changes that occur in the nucleosome and/or remodeler upon binding. Consistent with this idea, our SAXS analysis indicates that Chd1 binding in AMP-PNP and ADP-BeF_3 conditions promote distinct conformational states of nucleosomal DNA. Taken together, our findings suggest that conformational states of Chd1 coupled to the ATP binding/hydrolysis cycle stimulate dynamic unwrapping of the nucleosome. We propose that this Chd1-dependent unwrapping represents a new mechanism for remodeler altered histone-DNA interactions and DNA accessibility, which may elucidate the role of Chd1 in transcriptional elongation.

5.2 Materials and Methods

5.2.1 Protein expression and purification

The Chd1 proteins used in this study were from *Saccharomyces cerevisiae* and contained the chromodomains, ATPase motor, and DBD (residues 118-1274, referred

to throughout as Chd1) or only the chromodomains, ATPase motor, and the linker prior to the DBD (residues 118-1014, referred to as Chd1[ΔDBD]). Expression and purification of Chd1 proteins were carried out as previously described [16,19]. Histone proteins were from *Xenopus laevis*, and were expressed and purified as described [20]. For FRET experiments, a single cysteine was introduced (V35C) and labeled with Cy5-maleimide. Histones were prepared by refolding H2A and H2B into dimers and H3 and H4 into tetramers, purifying each by size exclusion chromatography.

5.2.2 DNA and nucleosome purification

Nucleosomal DNA was generated by large-scale PCR using the Widom 601 sequence [21] as the template, and purified by native acrylamide gel using a BioRad Prep Cell as previously described [20]. The DNA constructs are given in Supplementary Figure C.1. For FRET experiments, PCR reactions used DNA oligos containing Cy3 or Cy5 at the 5' end (IDT). Nucleosomes were reconstituted by using salt gradient dialysis to deposit purified histone octamer onto 601-containing DNA, mixed in a 1:1 ratio, as previously described [20].

5.2.3 FRET assays

Using fluorescently labeled nucleosomes, dyes were excited at 510 nm and emission spectra were collected from 530 nm to 700 nm on a Fluorolog 3 fluorometer (Horiba). For 16N16 nucleosomes, spectra were collected for 10 nM nucleosome first in the presence of 1X binding buffer (20 mM Hepes pH 7.5, 10 mM $MgCl_2$, 0.1 mM EDTA, 5% sucrose, 1 mM DTT, 0.02% Nonidet P-40, 0.1 mg/ml BSA, 100 mM KCl) and either

1 mM AMP-PNP or 1 mM ADP-BeF_3, and then again after addition of 30 nM Chd1. For 9N80 nucleosomes, spectra were collected for 20 nM nucleosomes in buffer plus nucleotide, and then separately for a mixture of nucleosome, nucleotide and Chd1. For these samples, nucleosomes were generated by addition of 32 nM H2A/H2B dimer (unlabeled) to 20 nM hexasomes. In ADP-BeF_3 conditions, 100 nM Chd1 or 1 μM Chd1(ΔDBD) was used. Due to the weak signal observed at lower concentrations, 1 μM Chd1 was also used under AMP-PNP conditions.

5.2.4 Stopped-flow experiments

The on-rates of Chd1 binding to nucleosomes were measured using an SX20 Stopped-Flow Spectrophotometer (Applied Photophysics). For these experiments, samples from two syringes were rapidly mixed. One syringe contained Chd1 at the specified concentrations, the other contained 20 nM 12N12 nucleosome. Both samples were suspended in 1X binding buffer (20 mM Hepes pH7.5, 10 mM $MgCl_2$, 0.1 mM EDTA, 5% sucrose, 1 mM DTT, 0.02% Nonidet P-40, 0.1 mg/ml BSA, 100 mM KCl without or with 1 mM AMP-PNP or ADP-BeF_3 [1 mM ADP, 1.2 mM $BeCl_2$, 6 mM NaF]). The nucleosome was Cy3-labeled on H4A15C. The dye was excited at 510 nm and the emission intensity collected using a 570 nm long-pass filter. For each set of syringes, progress curves from multiple injections were recorded at 25˚C. Typically 3-6 traces (technical replicates) were averaged together for each experiment. Fitting was carried out in Mathematica using the NonLinearModelFit function with the triple exponential form, $y = a_1\left(1 - e^{-k_1 x}\right) + a_2\left(1 - e^{-k_2 x}\right) + a_3\left(1 - e^{-k_3 x}\right) + c$. Here a_n were the amplitudes (fractions of total fluorescence range), k_n were the rates (sec^{-1}), and c was a

constant. The observed rates plotted versus Chd1 concentration provide the apparent k_{on} rates ($M^{-1} \cdot s^{-1}$), shown in Figure 5.2.

5.2.5 SAXS data collection and modeling

SAXS data were collected at G1 station at Cornell High Energy Synchrotron Source. The x-ray energy was 11.18 keV and sample-to-detector distance was 2.06 m (measured using a silver-behenate standard). Scattering intensities were collected on a Pilatus 200K detector over a q-range from ≈ 0.007-0.26 $Å^{-1}$. After azimuthal integration, SAXS profiles were normalized by the intensity of the primary beam, which was imaged directly on the detector, after attenuation by a 200 µM molybdenum beamstop. Samples were oscillated in a 2 mm quartz capillary with 10 µM thick walls. Multiple, 10 s exposures were collected and averaged. Profiles were carefully monitored for signs of radiation damage. 12N12 nucleosomes were incubated with Chd1 and nucleotides (ADP-BeF_3 or AMP-PNP) for at least 10 minutes before being measured. Buffers contained 10 mM Tris pH 7.8, 100 mM NaCl, 2 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 60% (w/v) sucrose, and ± 0.5 mM AMP-PNP or ADP-BeF_3 (0.5 mM ADP, 4 mM NaF, 0.6 mM $BeCl_2$). For each condition, 3-5 separate measurements were made and analyzed independently to confirm reproducibility.

Radius of gyration (Rg) values were calculated using GNOM [22]. To carry out ensemble modeling, a pool of 14,807 DNA structures was generated and theoretical SAXS profiles for each structure were calculated using CRYSOL [23] [1]. To select

[1] CRYSOL parameters: number of harmonics = 50, maximum s-value = 0.25, number of points = 201.

representative ensembles, we applied an ensemble optimization method [24–26] which exploits a genetic algorithm (GAJOE[2]) to optimize the selection of models from a pool whose summed scattering profiles best represents the experimentally measured profile, over the q range from 0.016-0.16 $Å^{-1}$. For the Chd1-12N12 complex measured in the presence of ADP-BeF_3 and AMP-PNP, the final best fitting ensembles for each of the 10 iterations of the genetic algorithm were pooled together and clustered using DAMCLUST [27].

5.3 Results

5.3.1 Chd1 binding reveals partial unwrapping by FRET

Reorganization of histone-DNA contacts within the nucleosome is believed to be principally driven by DNA translocation of the Chd1 ATPase motor, which requires turnover of ATP. However, recent work suggests that Chd1 binding unwraps DNA from the edge of the nucleosome [16,18]; we refer to the orientation of this unwrapped DNA as its trajectory. To monitor the influence of Chd1 in this process, we designed a FRET-labeled nucleosome having 16 bp of flanking DNA on each side of the strong Widom 601 positioning sequence (16N16), with Cy3 and Cy5 dyes at the DNA ends (Figure 1A). This labeling scheme yielded average FRET efficiencies of 0.45 ± 0.02 for the nucleosome alone, consistent with trajectories of entry/exit DNA of nucleosome crystal structures and indicative of a fully wrapped state. Addition of Chd1 in the

[2] GAJOE parameters: number of generations = 10,000, number of ensembles = 50, ensemble size fixed = no, maximum/minimum number of curves per ensemble = 20/5, curve repetition allowed: yes, constant subtraction = yes, number of times genetic algorithm repeated = 10.

presence of the ATP analogs AMP-PNP or ADP-BeF_3 markedly reduced FRET efficiencies to 0.088 ± 0.005 and 0.065 ± 0.015, respectively (Figure 5.1B). A decrease in FRET was also observed with Chd1 in the absence of nucleotide, but required the addition of considerably more Chd1 (1 μM for apo versus 30 nM for ATP analogs) to achieve a smaller change in FRET (0.45 to 0.32) (Supplementary Figure C.2).

This 16N16 nucleosome construct is inherently sensitive to DNA unwrapping due to the placement of the Cy3/Cy5 labels. By labeling both flanking DNA segments, FRET changes are amplified by the 2-fold symmetry of the nucleosome, as simultaneous unwrapping on both sides doubles the change in dye separation. Additionally, due to the superhelical geometry of the nucleosome, the DNA ends experience the greatest changes in position, increasing FRET sensitivity for small changes in the angle of entry/exit DNA. We also designed a different nucleosome with a Cy3/Cy5 FRET pair that would only respond to unwrapping on one side and would require a greater degree of unwrapping to separate the dye pair. Similar to the original nucleosome unwrapping experiments reported by Li and Widom [28], Cy3-maleimide was attached to histone H2A(T120C), and was therefore positioned adjacent to DNA entering/exiting nucleosome core, whereas a Cy5 label was attached to the DNA end on the 9 bp side of a 9N80 construct (Figure 5.1C). With this configuration, the nucleosome alone (20 nM) yielded FRET values of 0.342 ± 0.006. Upon addition of 100 nM Chd1 in the presence of ADP-BeF_3, the FRET value decreased to 0.164 ± 0.002 (Figure 5.1D). In AMP-PNP conditions, FRET also appeared to decrease, but the effect was more subtle. To ensure that the nucleosome was fully saturated with Chd1, we repeated experiments in AMP-PNP conditions with 1 μM Chd1, which also yielded a

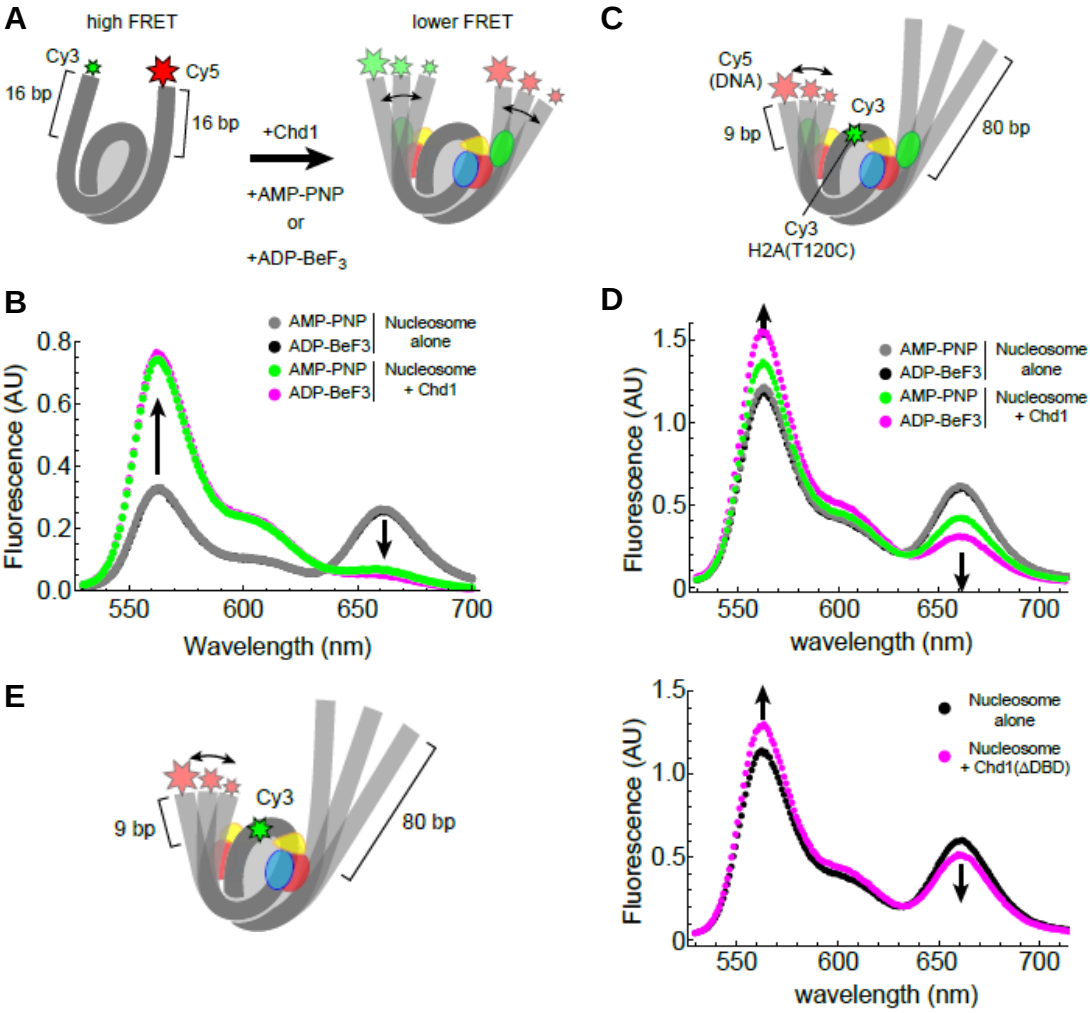


Figure 5.1 Chd1 binding promotes DNA unwrapping from the nucleosome**.**
(**A**) Schematic illustration of Cy3/Cy5 FRET pair positioning on the 16N16 nucleosome, with one dye on the 5' end of each DNA strand. Chd1 is colored by domains: chromodomains (yellow), ATPase motor (red/blue), and DNA-binding domain (green). (**B**) Wavelength emission scan for 10 nM 16N16 nucleosomes upon excitation at 510 nm (black, gray). As indicated by arrows, addition of 30 nm Chd1 promoted an increase in Cy3 emission (564 nm) and corresponding decrease in Cy5 emission (664 nm) in the presence of AMP-PNP (green) and ADP-BeF_3 (magenta). (**C**) Schematic illustration of 9N80 nucleosomes, with Cy5 on the DNA end of the 9 bp side, and Cy3 on histone H2A(T120C). (**D**) Wavelength emission scan for 20 nM 9N80 nucleosomes in the presence and absence of Chd1 upon excitation at 510 nm. Emission spectra for nucleosomes alone are shown in gray and black. ADP-BeF_3 conditions contained 100 nM Chd1 (magenta), whereas AMP-PNP conditions contained 1 µM Chd1 (green). (**E**) Schematic illustration of 9N80 nucleosomes with Chd1(ΔDBD) and wavelength emission scan for 20 nM 9N80 nucleosomes in the absence (black) and presence (magenta) of 1 µM Chd1(ΔDBD) upon excitation at 510 nm.

clear change in FRET to 0.239 ± 0.006, still smaller than the overall change observed with the addition of ADP-BeF_3 at lower concentration (Figure 5.1D). The magnitude of these FRET changes is consistent with the unwrapping of 10-20 bp of entry/exit DNA.

Because the Chd1 remodeler possesses several distinct domains, it is important to assess which domains are required for DNA unwrapping. We recently reported that the chromodomains and ATPase motor of Chd1 bind to internal sites on nucleosomal DNA, ~1 and ~2 helical turns from the nucleosome dyad, respectively, whereas the Chd1 DNA-binding domain (DBD) binds DNA immediately flanking the nucleosome on the opposite DNA gyre [16]. The chromo-ATPase is required to uniquely position the DBD at nucleosome edge, suggesting that these three domains formed a functional unit that spanned both DNA gyres of the nucleosome [16]. We suspected that a specific, nucleotide-dependent organization of Chd1 domains held the DBD in a position that favored an unwrapped trajectory of DNA exiting the nucleosome. To test this idea, we examined the extent that FRET changes could be stimulated by a Chd1 construct lacking the DBD. Unlike the construct containing chromo-ATPase and DBD, 30 nM Chd1(ΔDBD) failed to stimulate FRET changes in 10 nM nucleosome in the presence of ADP-BeF_3. However, a significant reduction in FRET (to 0.286 ± 0.006) was achieved with higher (1 μM) concentrations of Chd1(ΔDBD) (Figure 5.1E). These results are consistent with the idea that the DBD is required for high affinity binding, yet not necessary for stimulating changes in DNA trajectory. Instead, disruption of canonical nucleosomal wrapping is coupled to nucleotide-dependent conformational changes of the ATPase motor.

5.3.2 Nucleosome binding is associated with nucleotide-dependent conformational changes

Chd1 forms a high-affinity complex with the nucleosome in the presence of both AMP-PNP and ADP-BeF_3, yet these ATP analogs confer distinct properties to Chd1-nucleosome complexes. In the presence of ADP-BeF_3, nucleosome binding by Chd1 is less sensitive to the presence of flanking DNA and more resistant to increased ionic strength, with higher affinity Chd1-nucleosome complexes observed than with AMP-PNP (16). We hypothesize that these two nucleotide states favor distinct conformations of Chd1 on the nucleosome. To explore potential differences in these states, we measured the kinetics of nucleosome binding using stopped-flow. Previous work with the ISWI remodeler showed that nucleosome binding could be monitored by labeling the histone H4 tail (A15C) with Cy3; higher Cy3 fluorescence was observed for the bound state [29]. We discovered that for Chd1, Cy3 at this position increased binding affinity compared with unlabeled nucleosomes (Supplementary Figure C.3), and therefore previously reported measurements using H4(A15C-Cy3) nucleosomes likely overestimated affinities [16,19]. Interestingly, reactions performed with and without the Cy3 label typically required a substantially longer time (>20 min) to equilibrate in ADP-BeF_3 compared with AMP-PNP (<5 min), suggestive of slow conformational changes. We reasoned that the higher affinity toward H4(A15C-Cy3) nucleosomes should increase the on-rate or decrease the off-rate, and therefore rapid kinetics of binding could still be informative for detecting significant barriers in forming a stable Chd1-nucleosome complex.

To determine on-rates, rapid mixing experiments with 10 nM 12N12 H4(A15C-

Cy3) nucleosomes were carried out with increasing concentrations of Chd1 (Figure 5.2A). Binding reactions for both AMP-PNP and ADP-BeF_3 proceeded in several stages and were best fit with triple exponential equations. Plotting the observed rates with respect to concentration yielded apparent on-rates for both nucleotide conditions (Figure 5.2B). For ADP-BeF_3, all three rates were concentration dependent, whereas for AMP-PNP, only the fastest rate showed a clear concentration dependence. Interestingly, the fastest on-rates for AMP-PNP ($1.3 \pm 0.1 \times 10^7$ $M^{-1}s^{-1}$) and ADP-BeF_3 ($2.7 \pm 0.3 \times 10^6$ $M^{-1}s^{-1}$) were 10-100 times slower than expected for diffusion-limited binding. While the relatively slow binding observed for both nucleotide states likely reflects rate-limiting conformational changes required for stable interactions, the different binding characteristics suggests that upon formation of a complex, distinct conformations of the nucleosome and/or Chd1 are favored in each nucleotide state. Chd1 binding alters the path of nucleosomal DNA in distinct ways depending on the presence of ADP-BeF_3 or AMP-PNP.

The distinct binding characteristics associated with formation of Chd1-nucleosome complexes, shown in Figure 5.2, suggest that the nucleosome conformation may depend on the nucleotide state. To investigate this hypothesis, we applied contrast variation small angle x-ray scattering (CV-SAXS) to resolve the nucleotide-dependent conformation(s) of nucleosomal DNA changes upon Chd1 binding. SAXS provides insight into the global structures of individual macromolecules in solution; however, for large complexes containing multiple components with different electron densities (e.g. proteins and DNA), interpretation of SAXS data is limited due to the difficulty in

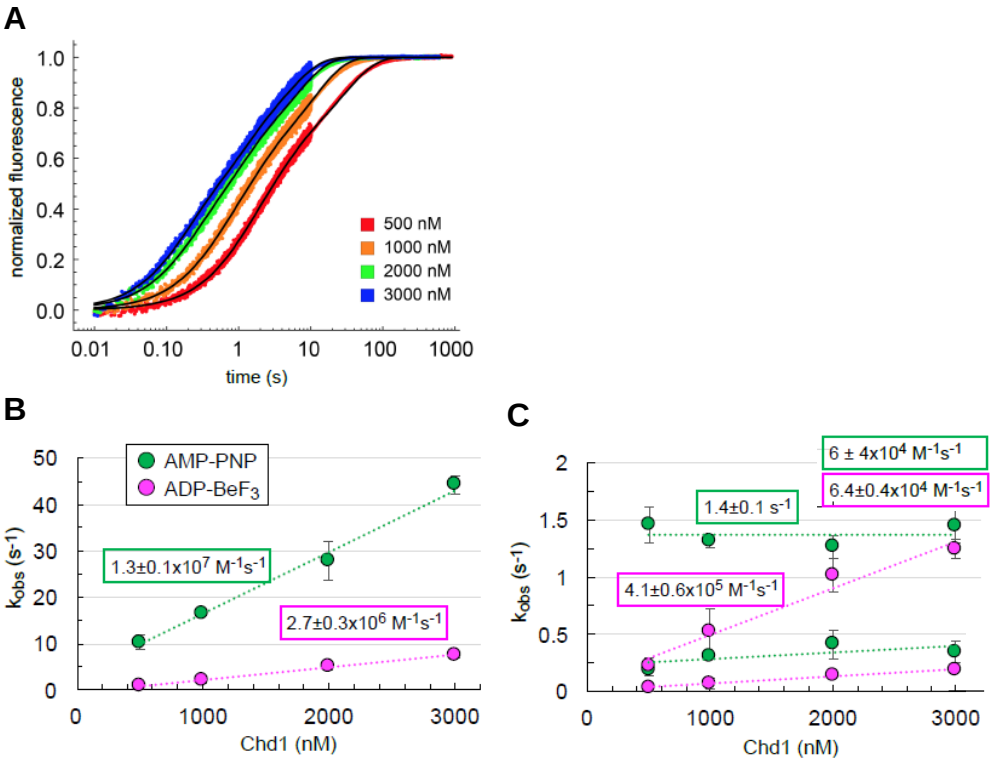


Figure 5.2 Nucleotide state affects the kinetics that Chd1 binds nucleosomes
(**A**) Progress curves from stopped-flow binding experiments of Chd1 with 10 nM 12N12 nucleosomes in the presence of ADP-BeF_3. Each curve is an average of 3-6 technical replicates. Shown is a representative titration of three independent experiments. (**B-C**) Plots of observed rate constants versus Chd1 concentration. For both AMP-PNP and ADP-BeF_3 conditions, the two slower rates (**C**) were plotted separately from the fastest rate (**B**) to more clearly show concentration dependence and rate differences. The calculated on-rates (slopes) show concentration dependence for all three observed rates in ADP-BeF_3 conditions, whereas under AMP-PNP conditions the middle rate failed to show a clear concentration dependence. Error bars show the standard deviation from three independent experiments.

resolving each component's contribution to the measured SAXS profile [30]. In CV-SAXS, sucrose is added to the bulk solvent until its electron density equals that of the lower electron density component, in this case the proteins. Under this matched condition, the protein components no longer contribute to the signal and details of the DNA conformation can be resolved. We previously used CV-SAXS to study DNA

unwrapping during the salt-induced disassembly of nucleosomes [31,32]. Here, we applied this technique to determine whether Chd1 altered the canonically wrapped conformation of nucleosomal DNA when bound in different nucleotide states.

The concentration of sucrose required to mask the protein signals was experimentally determined. Initially, SAXS profiles of the DNA (5 μM), Chd1 (15 μM), and histones (5 μM) were separately measured in solutions containing varying percentages of sucrose (Figure 5.3A). In 60% sucrose, the signals from protein components (Chd1 and histones) were effectively eliminated, but, due to its higher electron density, sufficient signal remained from the DNA to measure its conformation in the presence (or absence) of protein partners.

Figure 5.3B shows SAXS profiles measured in 60% sucrose for 12N12 nucleosomes (5 μM) with and without Chd1 (10 μM). All measured SAXS profiles are qualitatively similar, displaying the characteristic shape of DNA wound in nucleosomes, suggesting that most of the DNA remains wrapped in nucleosomal structures. As a control, we measured SAXS profiles for nucleosomes alone in the presence of different nucleotides, but without the remodeler. These profiles are identical (Supplementary Figure C.4A-B). Small but significant differences were detected when Chd1 was present, dependent on the presence of ADP-BeF_3 or AMP-PNP, or the absence of nucleotide (Apo) (Supplementary Figure C.4C-D). The radius of gyration (Rg) for the 12N12 nucleosome alone was 49.37 ± 0.26 Å, consistent with expectations for a fully wrapped structure. In the presence of Chd1, the Rg increased to 56.62 ± 0.73 Å and 55.26 ± 0.49 Å with the addition of ADP-BeF_3 and AMP-PNP, respectively, consistent with the partial unwrapping of the DNA ends observed using FRET (Figure 5.1B,D). The

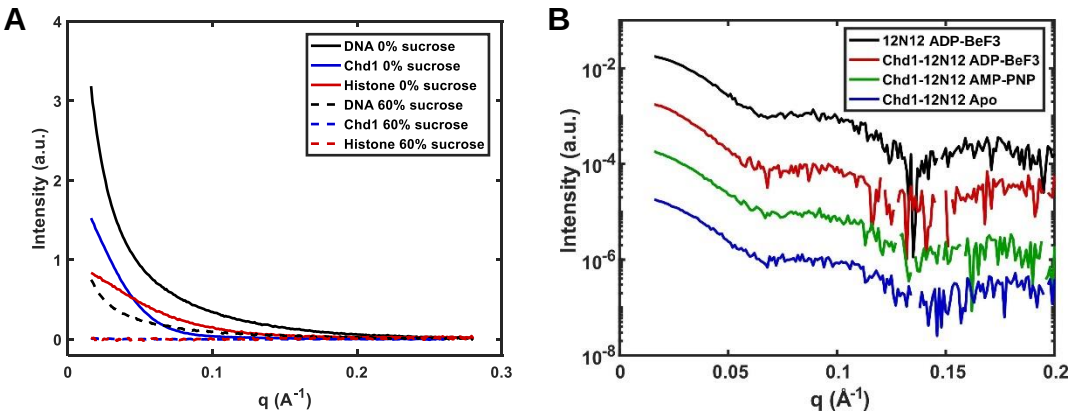


Figure 5.3 Protein contrast matching is achieved with 60% sucrose.
(**A**) SAXS signals are shown for DNA, Chd1, and histone octamer measured separately with 0% and 60% sucrose. In 0% sucrose, DNA (solid black), Chd1 (solid blue), histone (solid red) have distinct scattering profiles. In 60% sucrose, the DNA signal (dashed black) is reduced by a factor of ≈ 4, but scattering signals from Chd1 (dashed blue) and histone (dashed red) proteins vanish. (**B**) SAXS profiles measured in 60% sucrose for 12N12 in ADP-BeF_3, and Chd1-12N12 complex in ADP-BeF_3 (red), AMP-PNP (green), and Apo (blue).

complex (Chd1-12N12) in the apo state had an Rg of 50.35 ± 0.18 Å. This decreased response is consistent with the small changes observed in FRET for the apo state (Supplementary Figure C.2).

To gain structural insights that go beyond the average, measured Rg values, we applied an ensemble optimization method (EOM) to identify potential DNA structures present in the various complexes. Starting from a large pool of possible structures, EOM finds the subset of structures ("ensemble") that best recapitulates the SAXS data. Following the strategies in ref [31,32], we manually unwrapped DNA to generate a pool of structures. These structures were based on the nucleosome crystal structure 1KX5 [33] with varying amounts of DNA removed and replaced with straight segments (linear B-form DNA). This procedure produced 9,182 structures with the DNA unwrapped along the natural trajectory of the nucleosomal DNA. To further diversify the DNA

pool, we generated 5,625 variations of unwrapped nucleosomal DNAs that contained kinks (Supplementary Figure C.5). These kinks were generated by introducing various combinations of bends into 12N12 nucleosomal DNA using 3D-DART [34, 35]. From this pool of 14,807 total structures, a genetic algorithm [24–26] selected ensembles of DNA structures using the SAXS data collected for 12N12 nucleosomes with and without Chd1 in 60% sucrose. The resulting fits to the data are shown in Figure 5.4A-D.

To characterize the ensembles, we calculated Rg distributions for the structures selected by EOM (Figure 5.4E-F). Although the pool contained structures that ranged in size from 45-165 Å (Figure 5.4E), the structures within the selected ensembles cluster in size between 45-70 Å (Figure 5.4F), confirming that the DNA in 12N12 nucleosomes remains mostly wrapped in the presence of Chd1 regardless of the nucleotide state. The nucleosomes alone are represented by the peak centered around 49 Å. With ADP-BeF_3, Chd1-12N12 forms a prominent peak centered around 55 Å, reflecting partially unwrapped nucleosome structures. A similar, but slightly broader distribution is observed in the presence of AMP-PNP. In the apo (nucleotide-free) state, Chd1-12N12 appears to be a mixture between mostly wrapped nucleosomes and a small population of partially unwrapped structures. The EOM pools therefore suggest subtle but measurable differences in DNA conformation in response to nucleotide-dependent conformations of Chd1.

Representative ensembles for each condition are shown and compared in Figure 5.4G-J. The structures in each ensemble were assigned colors according to shared characteristic features. Under all conditions, at least some DNA remained mostly wrapped (blue in Figure 5.4G-J). These wrapped states are most prevalent for the

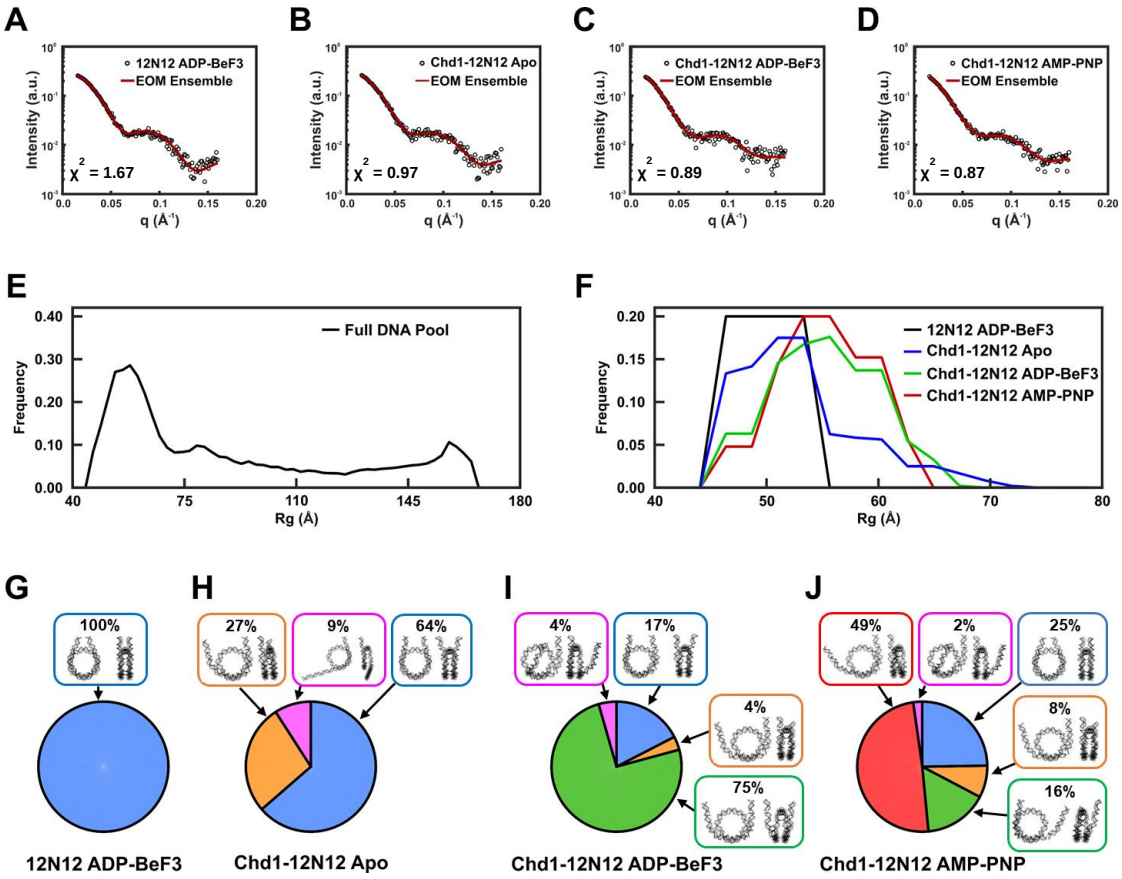


Figure 5.4 Representative ensembles for Chd1-12N12 in different nucleotide states. (**A-D**) SAXS data and EOM fits for nucleosome alone (12N12 ADP-BeF_3), and the Chd1-12N12 complex in the Apo, ADP-BeF_3, and AMP-PNP states, respectively. (**E-F**) Rg distributions for the DNA pool (E) and the selected ensembles (F) for nucleosome alone (12N12 ADP-BeF_3, black), and Chd1-12N12 complex in the Apo (blue), ADP-BeF_3 (green), and AMP-PNP (red) states. (**G-J**) Representative DNA structures selected by EOM for 12N12 and Chd1-12N12 in different nucleotide states. Structures are organized by color according to characteristic features as follows: blue = mostly wrapped, orange = in-plane unwrapping, green = out-of-plane unwrapping, red = in-plane wide unwrapping, magenta = minor populations.

nucleosomes without Chd1 (100%) and Chd1-12N12s in the apo state (64%). A quarter of the structures for Chd1-12N12 in AMP-PNP (25%) and a smaller fraction for Chd1-12N12 in ADP-BeF_3 (17%) remain mostly wrapped, consistent with the slightly stronger conformational response observed in FRET with ADP-BeF_3 compared with

AMP-PNP (Figure 5.1D).

The partially unwrapped structures observed in the presence of Chd1 show distinct asymmetric trajectories (orange, green, red in Figure 5.4G-J) that are present with varying weights depending on the nucleotide state. For all of the partially unwrapped structures (e.g. not the blue cluster), both DNA ends peel away from the histone at similar locations on both sides of the dyad axis. The shared origin of the unpeeling likely reflects common binding positions of Chd1 onto both sides of the nucleosome for all nucleotide states. However, in all cases, the unwrapping is more prominent on one side than the other. This asymmetry arises from varying angles in the trajectories of the DNA ends. Since SAXS is not sensitive to chirality, this analysis cannot distinguish unwrapping from a particular end of the DNA.

A notable difference between ADP-BeF_3 and AMP-PNP states was the extent of out-of-plane unwrapping. For Chd1-12N12 with ADP-BeF_3, the most prominent population (green, 75%) has both ends unwrapping out-of-plane ($\approx 20°$) relative to the plane perpendicular to the superhelical axis of nucleosomal DNA. In contrast, for Chd1-12N12 with AMP-PNP, the most prominent structure (red, 49%) has mostly in-plane unwrapping, with a significantly wider trajectory on one side. For Chd1-12N12 in the apo state, the most prominent unwrapped DNA structure (orange, 27%) is mostly in-plane but the DNA is not unwrapped to the same extent as in the presence of AMP-PNP. In summary, the two principal observations are as follows: (1) Chd1 binding leads to asymmetric unwrapping of the nucleosome ends, and (2) the trajectories of the unwrapped DNA depend on nucleotide-dependent conformational response of Chd1.

5.4 Discussion

This work demonstrates that Chd1 stimulates DNA unwrapping from the nucleosome in a nucleotide-dependent fashion. As for all ATPase motors, Chd1 is expected to be structurally sensitive to its nucleotide-bound state. Accordingly, we observed distinct kinetics of nucleosome binding in AMP-PNP versus ADP-BeF_3 conditions (Figure 5.2), which we interpret as reflecting conformational changes in Chd1 domains and/or the nucleosome upon binding. Remarkably, EOM analysis of CV-SAXS data strongly suggests that ADP-BeF_3 and AMP-PNP-bound states of Chd1 favor distinct unwrapped states of the nucleosome (Figures 5.4, 5.5). In AMP-PNP, the DNA unwraps at a wider angle but essentially in-plane with the nucleosome disk, whereas in ADP-BeF_3, the unwrapped DNA is more out-of-plane. These nucleotide-dependent differences suggest that DNA unwrapping is tightly coupled to motions of the ATPase motor. Previous modeling of Chd1 at SHL2 suggested that lobe 1 of the ATPase could directly contact entry/exit DNA of the opposite gyre [16], and here we show that the Chd1 DBD is not required for nucleosome unwrapping (Figure 5.1E). Taken together, these data support the idea that the Chd1 ATPase motor stimulates DNA unwrapping directly, with enhanced unwrapping dynamics during the ATP binding and hydrolysis cycle of Chd1.

A direct consequence of Chd1-induced DNA unwrapping from the nucleosome edge is the increased accessibility of DNA and histones, which in turn would be expected to alter binding of chromatin-associated factors. As observed in the first and many subsequent nucleosome crystal structures, residues 39-42 of histone H3 pass between the DNA gyres of the nucleosome at the DNA entry/exit site [36]. DNA unwrapping would expose this N-terminal segment of histone H3, and thus increase accessibility for

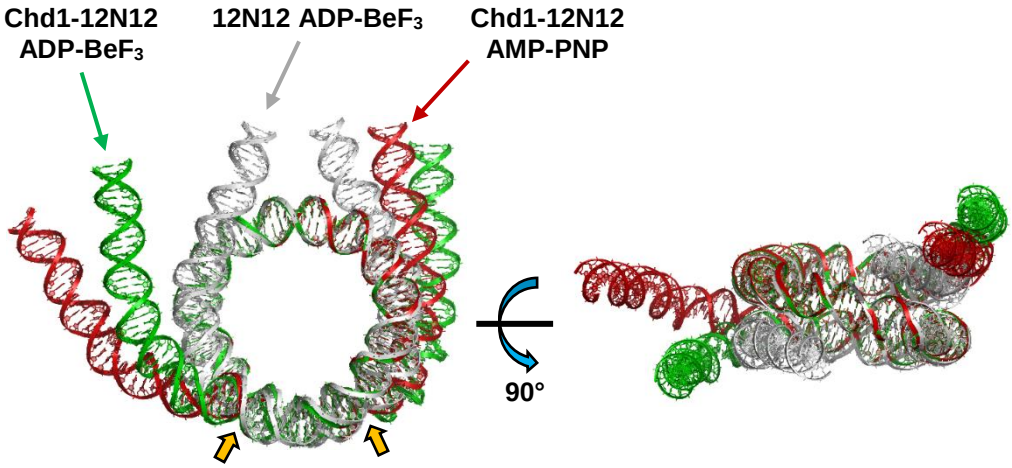


Figure 5.5 Comparison of most prominent DNA structures observed for Chd1-12N12 in ADP-BeF_3 and AMP-PNP, and 12N12 alone in ADP-BeF_3. Yellow arrows signify where the DNA begins to unwrap from the histone compared to the crystal structure (1KX5).

post-translational modifications such as H3R42 methylation and H3Y41 and H3T45 phosphorylation [37–41]. Methylation and acetylation of the nearby H3K36 is strongly coupled to transcription [42–44], and unwrapping may also favor deposition or recognition of these marks. In addition, unwrapping has been shown to increase binding of transcription factors to nucleosomal DNA, whose sites are otherwise poorly accessible [45,46]. Unwrapping is directly opposed by binding of linker histones H1 or H5, which stabilize the fully wrapped configuration of the nucleosome and have been found to reduce both transcription factor binding and phosphorylation of H3Y41 and H3T45 [39,40,47]. Interestingly, histone H1 is not incorporated by Chd1 during assembly [11], and the presence of H1 and H5 were found to also interfere with nucleosome sliding by Chd1 [48], suggesting an antagonistic relationship between Chd1 and linker histones.

Two defining characteristics of Chd1 interactions with nucleosomes are its abilities

to reposition and assemble nucleosomes, but it is unclear whether the DNA unwrapping described here is a key step for either of these processes. To our knowledge, DNA unwrapping by ISWI-type remodelers has not previously been described, yet ISWI remodelers have been shown to both assemble and slide nucleosomes in the presence of linker histones [11,48,49]. The action of ISWI remodelers on nucleosomes bound by linker histones argues against the idea that unwrapping is required for either nucleosome repositioning or assembly. Determining whether other remodeler types also weaken DNA-histone contacts at the entry/exit region of the nucleosome will be important for clarifying whether DNA unwrapping by Chd1 is unique or a shared characteristic among distinct remodeler families.

Separate from nucleosome repositioning and assembly activities, DNA unwrapping by Chd1 could provide a previously unappreciated means of altering nucleosome structure, which may be functionally coupled to transcription. Nucleosomes present a barrier to elongating polymerases such as RNA polymerase II (Pol II), causing transcriptional pausing [50,51]. Studitsky and coworkers have shown that DNA unwrapping from the upstream edge of the nucleosome facilitates transcription of Pol II through the outermost turn of nucleosomal DNA, whereas unwrapping from the downstream edge of the nucleosome greatly diminishes the major Pol II pause site at the upstream SHL2 [52]. Engagement of Chd1 with the downstream SHL2 would unwrap the upstream edge of the nucleosome and therefore potentially help Pol II entry. However, the presence of Pol II at the upstream SHL2, a major site of pausing, would sterically block Chd1 from unwrapping the downstream edge. Interestingly, Pol II remains paused at the upstream SHL2 site if the upstream DNA immediately behind the

polymerase does not rewrap [52]. DNA unwrapping, if maintained by Chd1 on the entry side, therefore would likely be antagonistic to Pol II transcribing through the nucleosome. Pol II transcription through nucleosomes is assisted by several elongation factors, and the FACT complex both reduces polymerase pausing and stabilizes histone-DNA contacts during transcription [53]. Although its mechanism of action is not presently clear, FACT appears to act in opposition to Chd1, with FACT mutant phenotypes in budding yeast rescued by Chd1 defects [54]. It has been shown that two polymerases traveling together evict nucleosomes [55], and therefore while speculative, controlled pausing may help keep Pol II complexes separated from each other. Our study raises the possibility that DNA unwrapping by Chd1 may contribute to transcription-related processes, providing an exciting new direction for future studies.

Acknowledgements

We thank Sarah Woodson for use of her stopped-flow and steady state fluorometers. Financial support for this work was provided by the National Institutes of Health (NIH) (R35- GM122514 to L.P.; R01-GM084192 and R01-GM113240 to G.D.B.; T32-GM007231 to R.J.T.). CHESS is supported by National Science Foundation (NSF) & National Institutes of Health/National Institute of General Medical Sciences (NIH/NIGMS) via NSF award [DMR-0936384 to CHESS]. Funding for open access charge: NIH R01-GM088645. We thank Arthur Woll for technical assistance at the G1 beamline.

REFERENCES

[1] R.K. McGinty, S. Tan, Recognition of the nucleosome by chromatin factors and enzymes, Curr. Opin. Struct. Biol. 37 (2016) 54–61.

[2] G.J. Narlikar, R. Sundaramoorthy, T. Owen-Hughes, Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes, Cell. 154 (2013) 490–503.

[3] A. Gaspar-Maia, A. Alajem, F. Polesso, R. Sridharan, M.J. Mason, A. Heidersbach, J. Ramalho-Santos, M.T. McManus, K. Plath, E. Meshorer, M. Ramalho-Santos, Chd1 regulates open chromatin and pluripotency of embryonic stem cells., Nature. 460 (2009) 863–868.

[4] M. Smolle, S. Venkatesh, M.M. Gogol, H. Li, Y. Zhang, L. Florens, M.P. Washburn, J.L. Workman, Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange, Nat. Struct. Mol. Biol. 19 (2012) 884–892.

[5] R. Simic, D.L. Lindstrom, H.G. Tran, K.L. Roinick, P.J. Costa, A.D. Johnson, G.A. Hartzog, K.M. Arndt, Chromatin remodeling protein Chd1 interacts with transcription elongation factors and localizes to transcribed genes, EMBO J. 22 (2003) 1846–1856.

[6] N.J. Krogan, M. Kim, S.H. Ahn, G. Zhong, M.S. Kobor, G. Cagney, A. Emili, A. Shilatifard, S. Buratowski, J.F. Greenblatt, RNA Polymerase II Elongation Factors of Saccharomyces cerevisiae: a Targeted Proteomics Approach, Society. 22 (2002) 6979–6992.

[7] D.E. Kelley, D.G. Stokes, R.P. Perry, CHD1 interacts with SSRP1 and depends on both its chromodomain and its ATPase/helicase-like domain for proper association with chromatin, Chromosoma. 108 (1999) 10–25.

[8] A.L. Hughes, O.J. Rando, Comparative Genomics Reveals Chd1 as a Determinant of Nucleosome Spacing in Vivo., G3 (Bethesda). 5 (2015) 1889–

1897.

[9] J. Ocampo, R. V. Chereji, P.R. Eriksson, D.J. Clark, The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo, Nucleic Acids Res. 44 (2016) 4625–4635.

[10] J. Fei, S.E. Torigoe, C.R. Brown, M.T. Khuong, G.A. Kassavetis, H. Boeger, J.T. Kadonaga, The prenucleosome, a stable conformational isomer of the nucleosome, Genes Dev. 29 (2015) 2563–2575.

[11] A. Lusser, D.L. Urwin, J.T. Kadonaga, Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly., Nat. Struct. Mol. Biol. 12 (2005) 160–166.

[12] R. Schwanbeck, H. Xiao, C. Wu, Spatial contacts and nucleosome step movements induced by the NURF chromatin remodeling complex, J. Biol. Chem. 279 (2004) 39933–39941.

[13] A. Saha, J. Wittmeyer, B.R. Cairns, Chromatin remodeling through directional DNA translocation from an internal nucleosomal site., Nat. Struct. Mol. Biol. 12 (2005) 747–755.

[14] M. Zofall, J. Persinger, S.R. Kassabov, B. Bartholomew, Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome, Nat. Struct. Mol. Biol. 13 (2006) 339–346.

[15] J.N. McKnight, K.R. Jenkins, I.M. Nodelman, T. Escobar, G.D. Bowman, Extranucleosomal DNA binding directs nucleosome sliding by Chd1., Mol. Cell. Biol. 31 (2011) 4746–59.

[16] I.M. Nodelman, F. Bleichert, A. Patel, R. Ren, K.C. Horvath, J.M. Berger, G.D. Bowman, Interdomain Communication of the Chd1 Chromatin Remodeler across the DNA Gyres of the Nucleosome, Mol. Cell. 65 (2017) 447–459.e6.

[17] V. Delmas, D.G. Stokes, R.P. Perry, A mammalian DNA-binding protein that

contains a chromodomain and an SNF2/SWI2-like helicase domain., Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 2414–8.

[18] R. Sundaramoorthy, A.L. Hughes, V. Singh, N. Wiechens, D.P. Ryan, H. El-Mkami, M. Petoukhov, D.I. Svergun, B. Treutlein, S. Quack, M. Fischer, J. Michaelis, B. Böttcher, D.G. Norman, T. Owen-Hughes, Structural reorganization of the chromatin remodeling enzyme Chd1 upon engagement with nucleosomes, Elife. 6 (2017) 1–28.

[19] I.M. Nodelman, K.C. Horvath, R.F. Levendosky, J. Winger, R. Ren, A. Patel, M. Li, M.D. Wang, E. Roberts, G.D. Bowman, The Chd1 chromatin remodeler can sense both entry and exit sides of the nucleosome, Nucleic Acids Res. (2016) gkw406.

[20] K. Luger, T.J. Rechsteiner, T.J. Richmond, Preparation of nucleosome core particle from recombinant histones, Methods Enzymol. 304 (1999) 3–19.

[21] P.T. Lowary, J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning., J. Mol. Biol. 276 (1998) 19–42.

[22] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.

[23] D. Svergun, C. Barberato, M.H. Koch, CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 28 (1995) 768–773.

[24] P. Bernadó, E. Mylonas, M. V Petoukhov, M. Blackledge, D.I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering., J. Am. Chem. Soc. 129 (2007) 5656–64.

[25] P. Bernadó, D.I. Svergun, Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering, Mol. BioSyst. 8 (2012) 151–167.

[26] G. Tria, H.D.T. Mertens, M. Kachala, D.I. Svergun, Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering, IUCrJ. 2 (2015) 207–217.

[27] M. V. Petoukhov, D. Franke, A. V. Shkumatov, G. Tria, A.G. Kikhney, M. Gajda, C. Gorba, H.D.T. Mertens, P. V. Konarev, D.I. Svergun, New developments in the ATSAS program package for small-angle scattering data analysis, J. Appl. Crystallogr. 45 (2012) 342–350.

[28] G. Li, J. Widom, Nucleosomes facilitate their own invasion., Nat. Struct. Mol. Biol. 11 (2004) 763–769.

[29] J.D. Leonard, G.J. Narlikar, A Nucleotide-Driven Switch Regulates Flanking DNA Length Sensing by a Dimeric Chromatin Remodeler, Mol. Cell. 57 (2015) 850–859.

[30] J.M. Tokuda, S.A. Pabit, L. Pollack, Protein-DNA and ion-DNA interactions revealed through contrast variation SAXS, Biophys. Rev. 8 (2016) 139–149.

[31] Y. Chen, J.M. Tokuda, T. Topping, J.L. Sutton, S.P. Meisburger, S. A. Pabit, L.M. Gloss, L. Pollack, Revealing transient structures of nucleosomes as DNA unwinds, Nucleic Acids Res. 42 (2014) 8767–8776.

[32] Y. Chen, J.M. Tokuda, T. Topping, S.P. Meisburger, S.A. Pabit, L.M. Gloss, L. Pollack, Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core, Proc. Natl. Acad. Sci. 114 (2017) 334–339.

[33] T.J. Richmond, C.A. Davey, The structure of DNA in the nucleosome core., Nature. 423 (2003) 145–50.

[34] M. van Dijk, A.M.J.J. Bonvin, 3D-DART: A DNA structure modelling server, Nucleic Acids Res. 37 (2009) 235–239.

[35] X.J. Lu, W.K. Olson, 3DNA: A software package for the analysis, rebuilding

and visualization of three-dimensional nucleic acid structures, Nucleic Acids Res. 31 (2003) 5108–5121.

[36] K. Luger, A.W. Mäder, R.K. Richmond, D.F. Sargent, T.J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution., Nature. 389 (1997) 251–260.

[37] F. Casadio, X. Lu, S.B. Pollock, G. LeRoy, B. a Garcia, T.W. Muir, R.G. Roeder, C.D. Allis, H3R42me2a is a histone modification with positive transcriptional effects., Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 14894–9.

[38] S.P. Baker, J. Phillips, S. Anderson, Q. Qiu, J. Shabanowitz, M.M. Smith, J.R. Yates, D.F. Hunt, P.A. Grant, Histone H3 Thr 45 phosphorylation is a replication-associated post-translational modification in S. cerevisiae., Nat. Cell Biol. 12 (2010) 294–8.

[39] M.A. Dawson, A.J. Bannister, B. Göttgens, S.D. Foster, T. Bartke, A.R. Green, T. Kouzarides, JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin., Nature. 461 (2009) 819–22.

[40] S.M. Jang, S. Azebi, G. Soubigou, C. Muchardt, DYRK1A phoshorylates histone H3 to differentially regulate the binding of HP1 isoforms and antagonize HP1-mediated transcriptional repression., EMBO Rep. 15 (2014) 686–94.

[41] M. Brehove, T. Wang, J. North, Y. Luo, S.J. Dreher, J.C. Shimko, J.J. Ottesen, K. Luger, M.G. Poirier, Histone core phosphorylation regulates DNA accessibility, J. Biol. Chem. 290 (2015) 22612–22621.

[42] T. Xiao, H. Hall, K.O. Kizer, Y. Shibata, M.C. Hall, C.H. Borchers, B.D. Strahl, Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast, Genes Dev. 17 (2003) 654–663.

[43] O. Bell, C. Wirbelauer, M. Hild, A.N.D. Scharf, M. Schwaiger, D.M. MacAlpine, F. Zilbermann, F. van Leeuwen, S.P. Bell, A. Imhof, D. Garza,

A.H.F.M. Peters, D. Schübeler, Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila., EMBO J. 26 (2007) 4974–84.

[44] S.A. Morris, B. Rao, B.A. Garcia, S.B. Hake, R.L. Diaz, J. Shabanowitz, D.F. Hunt, C.D. Allis, J.D. Lieb, B.D. Strahl, Identification of histone H3 lysine 36 acetylation as a highly conserved histone modification, J. Biol. Chem. 282 (2007) 7632–7640.

[45] K.J. Polach, J. Widom, Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation., J. Mol. Biol. 254 (1995) 130–49.

[46] J.A. North, J.C. Shimko, S. Javaid, A.M. Mooney, M.A. Shoffner, S.D. Rose, R. Bundschuh, R. Fishel, J.J. Ottesen, M.G. Poirier, Regulation of the nucleosome unwrapping rate controls DNA accessibility, Nucleic Acids Res. 40 (2012) 10215–10227.

[47] M. Bernier, Y. Luo, K.C. Nwokelo, M. Goodwin, S.J. Dreher, P. Zhang, M.R. Parthun, Y. Fondufe-Mittendorf, J.J. Ottesen, M.G. Poirier, Linker histone H1 and H3K56 acetylation are antagonistic regulators of nucleosome dynamics, Nat. Commun. 6 (2015) 10152.

[48] V.K. Maier, M. Chioda, P.B. Becker, ATP-dependent chromatosome remodeling, Biol. Chem. 389 (2008) 345–352.

[49] B.-R. Zhou, H. Feng, H. Kato, L. Dai, Y. Yang, Y. Zhou, Y. Bai, Structural insights into the histone H1-nucleosome complex, Proc Natl Acad Sci. 48110 (2013) 19390–19395.

[50] M.L. Kireeva, W. Walter, V. Tchernajenko, V. Bondarenko, M. Kashlev, V.M. Studitsky, Nucleosome Remodeling Induced by RNA Polymerase II, Mol. Cell. 9 (2002) 541–552.

[51] M.L. Kireeva, B. Hancock, G.H. Cremona, W. Walter, V.M. Studitsky, M.

Kashlev, Nature of the nucleosomal barrier to RNA polymerase II, Mol. Cell. 18 (2005) 97–108.

[52] O.I. Kulaeva, D.A. Gaykalova, N.A. Pestov, V. V Golovastov, D.G. Vassylyev, I. Artsimovitch, V.M. Studitsky, Mechanism of chromatin remodeling and recovery during passage of RNA polymerase II, Nat Struct Mol Biol. 16 (2009) 1272–1278.

[53] F.-K. Hsieh, O.I. Kulaeva, S.S. Patel, P.N. Dyer, K. Luger, D. Reinberg, V.M. Studitsky, Histone chaperone FACT action during transcription through chromatin by RNA polymerase II., Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 7654–9.

[54] D. Biswas, R. Dutta-Biswas, D.J. Stillman, Chd1 and yFACT act in opposition in regulating transcription., Mol. Cell. Biol. 27 (2007) 6279–87.

[55] O.I. Kulaeva, F.-K. Hsieh, V.M. Studitsky, RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones., Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 11325–30.

Chapter 6

Conclusions

In this work, we used small angle x-ray scattering to investigate partially unwrapped nucleosome conformations. Interpretation of SAXS data is complicated by two factors: (1) the difficulty in resolving how the protein and DNA components contribute to the measured signal, and (2) the conformational averaging that arises from the ensemble nature of the measurement. We addressed the first issue by using contrast variation methods to mask the protein's contribution to the scattering signal. The resulting scattering comes from the DNA and enables targeted studies of DNA conformations within the complex. To resolve structural details beyond the conformational average, we applied ensemble optimization methods and determined ensembles of DNA structures that best recapitulate the measured SAXS profiles. These strategies provided new opportunities to visualize how nucleosomal DNA unwraps in response to different conditions.

In Chapters 3-4, we used salt to globally destabilize histone-DNA interactions and populate partially disrupted nucleosomes [1,2]. To monitor the dynamics and get a better understanding of how nucleosomes unravel to permit DNA access, we used a stopped flow mixer to induce salt jumps. Above 0.5 M NaCl, we observed that DNA unwraps from the ends and found that the extent and rate of unwrapping increased as a function of salt concentration. At high salt concentrations (2.0 M NaCl), where full disassembly is achieved, we found that the DNA rapidly unwraps, but forms an asymmetric intermediate that persists for several hundred milliseconds. At intermediate

salt concentrations (1.2 M NaCl), where only partial disassembly is expected, we resolved two pathways through which the asymmetric intermediate is formed. In light of the symmetric nature of the nucleosome structure, the asymmetric unwrapping is surprising. Asymmetric flexibility between the two sides of the DNA sequence has been proposed as a mechanism that imparts this asymmetry [3]. This hypothesis can be tested by repeating these experiments with palindromic DNA sequences. FRET labels incorporated within the nucleosome were used to monitor the release of histone proteins. At the intermediate salt concentration, asymmetric DNA unwrapping appeared to coordinate the asymmetric release of the histone proteins. This suggests a new mechanism through which the exchange of histone proteins can be facilitated.

In Chapter 5, we investigated the conformational response of nucleosomes after being bound by the chromatin remodeler Chd1. Both FRET and SAXS reveal partial DNA unwrapping in the presence of Chd1. Although the trajectories of the unwrapped DNA vary significantly from those observed in Chapters 3-4 using salt, the unwrapped DNA in the presence of Chd1 also appeared to be asymmetric. This suggests that the asymmetric response between the two sides is an inherent property of the nucleosome. Interestingly, we found that Chd1 promotes distinct conformational changes depending on the nucleotides present (ADP-BeF_3, AMP-PNP, or apo). This observation implicates the ATPase motor as a key element in disrupting the histone-DNA contacts. The asymmetrically unwrapped nucleosomes observed in the presence of Chd1 may serve as an important functional configuration helps RNA polymerase II gain entry into the edge of the nucleosome. The inherent propensity for nucleosomes to unwrap asymmetrically may serve as an important mechanism to promote the proper directional

engagement of nuclear machinery (e.g. 5' to 3' on the sense strand). Features that bias asymmetric unwrapping can be encoded into the DNA sequence.

The methods applied in this work can be easily adapted to provide structural insight other nucleosomal interactions as well as the influence of different nucleosome modifications. The conformational activity of transcription factors, histone chaperones, and other chromatin remodelers can be examined. Nucleosomes containing various DNA sequences, histone variants, and post-translational modifications may exhibit new behaviors and reveal new mechanisms through which nucleosomes regulate genetic activity.

Developments in instrumentation has great potential to expand the scope and accessibility of the technique. One of the primary challenges for this approach is the difficulty in preparing consistent samples due to the high viscosity of the solutions. Advancements in mixing and sample manipulation technologies would greatly lower the technical barrier and most importantly the frustration level of many researchers. The development of continuous flow mixers compatible with high viscosity solutions would enhance the data quality collected for short timescales, and has the potential to reduce sample requirements. Since the ensemble method is wholly reliant on the quality of the initial DNA pool, advancements in structural modeling that integrates biophysical and biochemical data has great potential to improve the quality of the results.

The general strategy applied in this work is a powerful platform that can be applied to study virtually any protein-nucleic acid complexes. RNA-protein systems drive many of the most fundamental processes inside the cell. This technique can be used to provide insights into the activity and folding of RNA-protein complexes. The

organization and processing of DNA/RNA in viruses and bacteriophage may also be an interesting target. Advancements in high-throughput genomics are revealing new connections between genes and their products on a pace that greatly dwarfs structural determination – providing exciting opportunities to use SAXS to uncover the structure-function relationships within chromatin that coordinate biological networks!

Appendix A

Supporting Information for Chapter 3

Revealing transient structures of nucleosomes as DNA unwinds

A.1 Supplementary Materials and Methods

A.1.1 Histone Purification

The *Xenopus laevis* histones were expressed recombinantly as inclusion bodies in *E. coli* using previously described T7-expression plasmids [1]. Purification of the individual histones and reconstitution to H2A-H2B and $(H3\text{-}H4)_2$ oligomers was performed as described previously [2,3], except for the final chromatography step over a heparin column, described previously for H2A-H2B purification. First, this step was added to the purification of $(H3\text{-}H4)_2$ as well. Secondly, the chromatography buffers were altered to Tris and NaCl (rather than potassium phosphate and KCl) so that the histones were eluted in similar conditions to those employed in NCP reconstitution. Protein concentrations were determined by absorbance at 280 nm with extinction coefficients of 10,240 and 17,920 $M^{-1}cm^{-1}$ for the H2A-H2B dimer and $(H3\text{-}H4)_2$ tetramer, respectively. The [NaCl] was determined by refractive index.

A.1.2 DNA Production

A plasmid with 24 tandem repeats of the central 149 bp of the Widom 601 DNA was described previously [4], constructed using the iterative approach of ref. [5]. For consistency in comparing SAXS data with different DNA sequences, a similar plasmid

was constructed for the nucleosome positioning element of the *Lytechinus variegatus* 5S rRNA gene and promoter. The central 144 bp positioning element was PCR-amplified from the pSL208-12 plasmid (a gift from Dr. Michael J. Smerdon's lab, Washington State University) with primers that introduced EcoRV restriction sites at the ends of the PCR product. This fragment was used to engineer a plasmid with 24 tandem repeats. Subsequent plasmid purification, EcoRV digestion to liberate 149 bp fragments and their purification was carried out by protocols detailed elsewhere [4,5].

601-DNA sequence: [ATCGG AGAAT CCCGG TGCCG AGGCC GCTCA ATTGG TCGTA GACAG CTCTA GCACC GCTTA AACGC ACGTA CGCGC TGTCC CCCGC GTTTT AACCG CCAAG GGGAT TACTC CCTAG TCTCC AGGCA CGTGT CAGAT ATATA CATCC CGAT].

5S-DNA sequence: [ATCTC CAGGG ATTTA TAAGC CGATG ACGTC ATAAC ATCCC TGACC CTTTA AATAG CTTAA CTTTC ATCAA GCAAG AGCCT ACGAC CATAC CATGC TGAAT ATACC GGTTC TCGTC CGATC ACCGA AGTCA AGCAG CATAG GGCTC GGAT].

A.1.3 Contrast Variation

For SAXS, x-ray photons scatter off particles in solution with excess electron density $\Delta\rho_p = \rho_p - \rho_s$, called contrast, where ρ_p and ρ_s are the average electron densities of the particle and solvent, respectively. This contrast generates a scattering pattern that is distinct from the uniform background scattering of the buffer. For a biomacromolecule, both the overall shape and internal structure contributes to the scattering. However, both the protein and nucleic acid in a complex can be treated as homogeneous components

since the difference in average electron densities between the two components is much greater than the electron density fluctuations within each component. Thus, the scattering intensity as a function of contrast for a protein-nucleic acid (PNA) complex can be simplified as [6]

$$I(q) = (\Delta\rho_1)^2 I_1(q) + 2\Delta\rho_1\Delta\rho_2 I_{12}(q) + (\Delta\rho_2)^2 I_2(q)\,, \qquad \text{(Supplementary Eq. 1)}$$

where $\Delta\rho_1$,$I_1(q)$ and $\Delta\rho_2$, $I_2(q)$ represent the contrasts and scattering intensities of the protein and nucleic acid, respectively. $I_{12}(q)$ is the cross-term intensity.

By increasing the solvent electron density ρ_s to match that of the protein ρ_1, the first and second terms in Supplementary Eq. 1 vanish and the resulting scattering profile is dominated by the nucleic acid component (the third term). Consequently, scattering data containing information about the conformation of nucleic acids in protein complexes can be isolated. The addition of 50% (w/v) sucrose was found to be sufficient to blank histone proteins. This concentration of sucrose had no effect on the NaCl-dependent equilibrium stability of the NCP, as monitored by a previously described FRET system[6]. The following table shows the average electron densities for biomolecules and the solutions used [6]. Note that additional molecules (e.g. salts, buffers) in solution also contribute to the actual values.

Average electron densities	
Component	ρ (e^-/nm^3)
Water	334
50% Sucrose (w/v)	400
Proteins	420
Nucleic Acids	550

The contrast dependence of any parameters measured in SAXS must be taken into account. For SAXS experiments with sufficiently high contrast, e.g. when using buffers without sucrose or high salt concentrations, these effects are typically negligible.

A.1.4 I(0) Analysis

In SAXS, the average mass of the scattering particles is related to the extrapolated scattering intensity at zero scattering angle I(q = 0). Interpretation of I(0) requires knowledge of both sample heterogeneity and contrast. It is important to note that the contrast depends not only on sucrose, but also on [NaCl]. Since a wide range of [NaCl] was used for the equilibrium experiments, we limited our analysis to the endpoints with the assumption of monodispersity (fully associated octamers in 0.2 M NaCl and fully dissociated in 2.0 M NaCl). The static and kinetic data were measured at the same sample concentrations, allowing for relative comparison of their I(0)s (Figure 3.5A,B). To take into account the contrast difference between 0.2 M and 1.9 M [NaCl], we applied a scaling factor for the 0.2 M NaCl I(0). This factor was determined through CRYSOL [7] by calculating theoretical I(0)s at two different solvent electron densities for the same wrapped structure (2 M NaCl adds ≈ 30 e^-/nm^3 to pure water). Since time dependent changes using a stopped flow mixer are measured against a fixed background, the contrast does not change with time. Consequently, changes in I(0) report changes in the molecular mass and association state of the proteins as the nucleosomes disassemble. Because multiple association states may be present, we restricted our analysis to the endpoints— which agreed well with the equilibrium results

(Figure 3.5A,B).

A.1.5 Radius of gyration analysis

The radius of gyration (Rg) reported in SAXS is the root mean square distance between the scattering electrons within the particle. The apparent Rg and its contrast dependence is represented as [6]

$$Rg^2 = {R_c}^2 + \frac{\alpha}{\Delta\rho} - \frac{\beta}{(\Delta\rho)^2} \ , \qquad \text{(Supplementary Eq. 2)}$$

where R_c is the radius of gyration observed for particles at infinite contrast, $\Delta\rho$ is the contrast, and α and β are parameters that reflect the electron density fluctuations for particles in solution. α characterizes the internal structure of a particle. For a core-shell particle, α is positive when the shell has higher electron density than the core, and negative in the opposite case. β is always positive and increases as the apparent center of mass for one component is displaced relative to the other as a function of contrast.

Rgs reported in this work were determined using GNOM since the analysis uses the entire scattering curve and is less susceptible to low-q noise (in contrast to Guinier analysis). The contrast dependent contributions to Rg (second and third terms in Supplementary Eq. 2) were found to be negligible for the range of contrasts involved in this study. Theoretical Rgs determined from NCP models with the solvent electron densities varied from 334-430 e^-/nm^3 showed less than ± 1 Å deviations to the apparent Rgs.

A.1.6 Singular Value Decomposition

Singular value decomposition (SVD) is a powerful strategy for determining the minimum number of components necessary to reconstruct a dataset collected over a

series of experimental conditions (e.g. over time or a range of NaCl conditions). A set of data plotted with *m* discrete momentum transfer values and *n* scattering curves is arranged as an *m x n* matrix *A*, where each column represents the scattering profile collected within an experimental series. An SVD algorithm (MATLAB) identifies orthogonal basis curves by representing the matrix as a product of three matrices,

$$A = UWV^T \ . \qquad \text{(Supplementary Eq. 3)}$$

U is an *m x n* matrix of orthogonal columns that form a complete set of basis curves through which the entire series of scattering profiles can be represented by linear superposition. *W* is an *n x n* diagonal matrix containing the *singular values* conventionally ordered by decreasing value. Each singular value w_j represents the overall weight of the basis component U_j. V^T is the transpose of an *n x n* matrix *V*, where columns V_j represents the dependence of columns U_j on the course of the series. Therefore, each column in WV^T contains the linear superposition coefficients for the basis curves at each point in the series. The basis components that make significant signal contributions can be determined by direct inspection and by comparing the corresponding singular values. Although the number of distinctly scattering species corresponds with the number of basis components, the basis components themselves do not necessarily represent morphological SAXS profiles. Independent SAXS curves are linear combinations of the basis components and can be approximated using the independent basis curves U_j(q) as

$$I(q) = \sum_{j=1}^{r} U_j(q) w_j V_j \ , \qquad \text{(Supplementary Eq. 4)}$$

where *r* is the effective rank, or number of independent basis curves necessary to reconstruct the dataset.

A.1.7 Mixer Characterization

Mixing was characterized through fluorescence assays using N-acetyltryptophanamide (NATA) fluorescence quenching by N-bromosuccinimide (NBS) and Alexa350 quenching by KI. The mixing dead time was measured to be ≈ 6.6 ms using NATA-NBS at a flow rate of 6 mL/s. Incorporation of the HDS mixer reduced convection and back flow, allowing access to timescales up to 60 s. The optimal flow rates and volumes used were 6 mL/s and 315 µL for 0% sucrose and 7.5 mL/s and 375 µL for 50% sucrose. For SAXS experiments, complete salt mixing was confirmed by monitoring the buffer scattering. Mixing between subsequent shots was found to be consistent.

A.1.8 Minimum chi-square ($\overline{\chi^2}$) fit

Agreement between the experimentally observed 601-NCP kinetic intermediate and potential conformational models (Supplementary Figure A.8) was first assessed by evaluating the following chi-square:

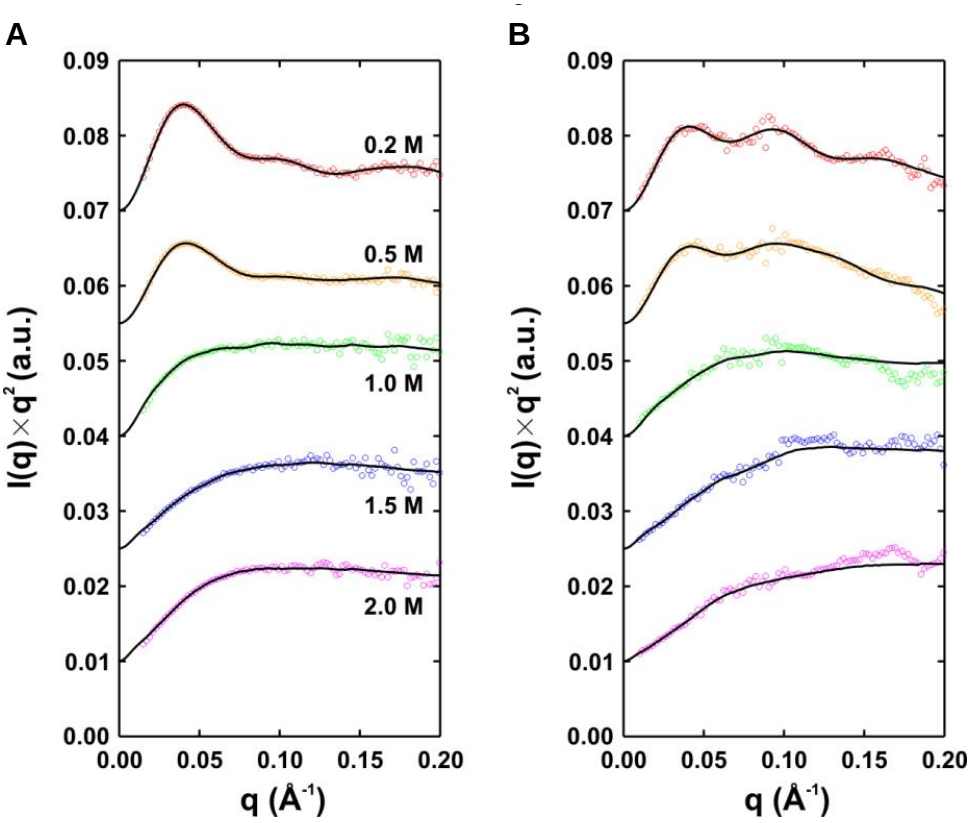
$$\overline{\chi^2} = \frac{1}{M}\sum_{i=1}^{M}\left(\frac{I_{\mathrm{exp}}(q_i) - I_{\mathrm{model}}(q_i)}{\sigma_{\mathrm{exp}}(q_i)}\right)^2, \qquad \text{(Supplementary Eq. 5)}$$

where $I_{\mathrm{exp}}(q_i)$ is the experimental scattering intensity at q_i, $I_{\mathrm{model}}(q_i)$ is the scattering intensity calculated from PDB models using CRYSOL, $\sigma_{\mathrm{exp}}(q_i)$ is the experimental error and M is the number of data points in q space. The best fit is revealed by a minimal $\overline{\chi^2}$ value, where a $\overline{\chi^2}$ value less than 1 indicates good agreement. The q-range used was limited to 0.015-0.12 $Å^{-1}$ to reduce the effects of high-q noise.

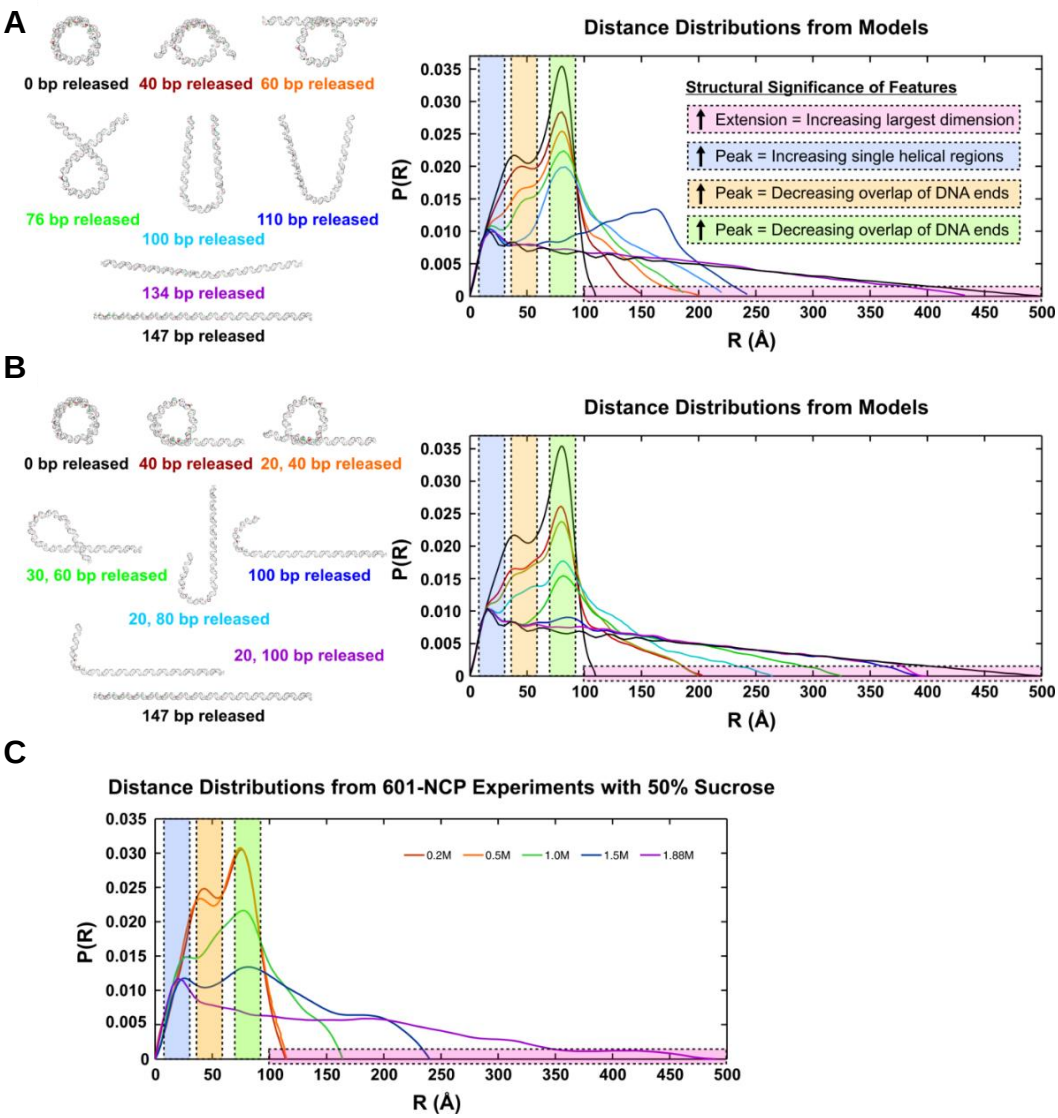
A.1.9 Ensemble Optimization Method (EOM)

To further investigate possible polydispersity and structural fluctuation, we applied the program GAJOE (Genetic Algorithm Judging Optimization of Ensembles) to determine an ensemble of DNA conformational models, whose combined theoretical scattering intensity best describes the experimental SAXS data (e.g. that of the 601-NCP kinetic intermediate). GAJOE uses a genetic algorithm described in ref. [8]. The q-range for GAJOE fitting was also limited to 0.015-0.12 $Å^{-1}$ to reduce the effects of high-q noise. When comparing to the 601-NCP intermediate (observed in the first 200 ms), one optimized ensemble was first generated from a pool of DNA PDB models (Group 1, 2 and 3 in Supplementary Figure A.8), and was found to be populated with mostly asymmetric models. Hence, 20 more asymmetric models (group 4) structurally similar to the most picked model in the first round (the 'Long 80bp released, Short 20bp released' model) were added to the pool, and a new process of ensemble optimization was conducted. The final optimized ensemble contains only similar asymmetric models (Supplementary Figure A.8 and Figure 3.6E), indicative of a nearly homogenous intermediary state for the 601-NCP. The $\overline{\chi^2}$ value of 0.90 is also lower than that minimized from single model $\overline{\chi^2}$ fit (0.95 for the 'Long 80bp released, Short 20bp released' model), indicating a better agreement to the experimental measurement when ensemble optimization method is used.

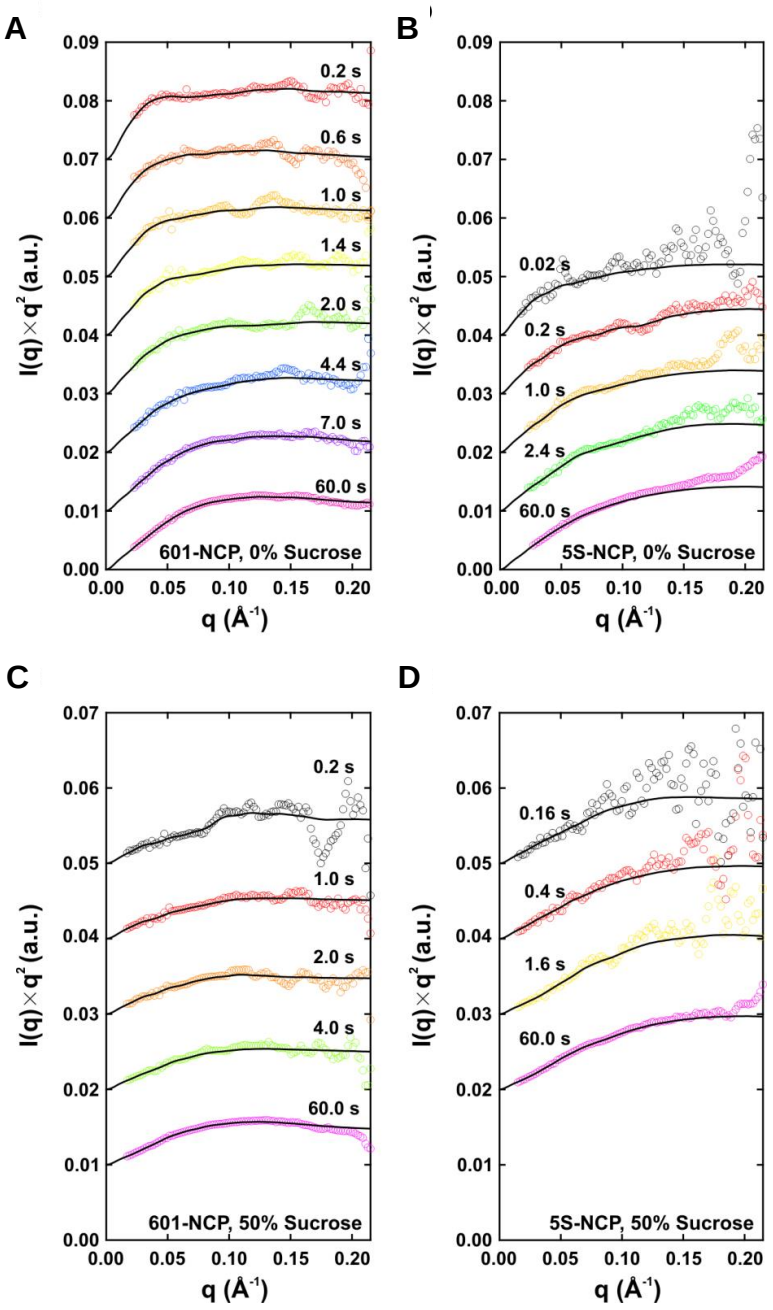
A.2 Supplementary Figures



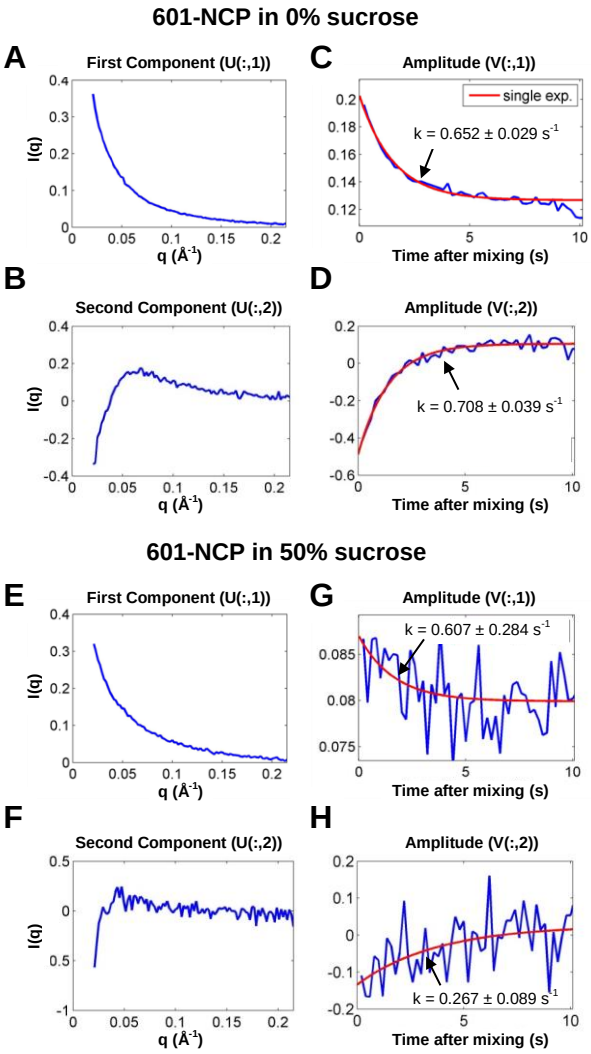
Supplementary Figure A.1 Kratky plots for 5S-NCP at different [NaCl] with (**A**) 0% and (**B**) 50% sucrose. The data (colored circles) and regularized fits to the data using GNOM (black lines) were scaled and offset to enhance visualization. A general transition from a globular to an extended structure is observed as the NaCl concentration was increased from 0.2 M to 2.0 M. Many more structural details (visualized as peaks and troughs) are distinguishable with the proteins blanked in (B) 50% sucrose compared to the (A) 0% sucrose data. Interestingly, the peaks for the 5S-NCP in low NaCl were generally more broadened compared to the 601-NCP (Figure 3.3A,B), reflecting greater conformational variation. Furthermore, the 5S-NCP appears mostly unwrapped at 1 M NaCl, suggesting that 5S-NCP is less stable than 601-NCP. Plots were scaled and offset to enhance visualization.



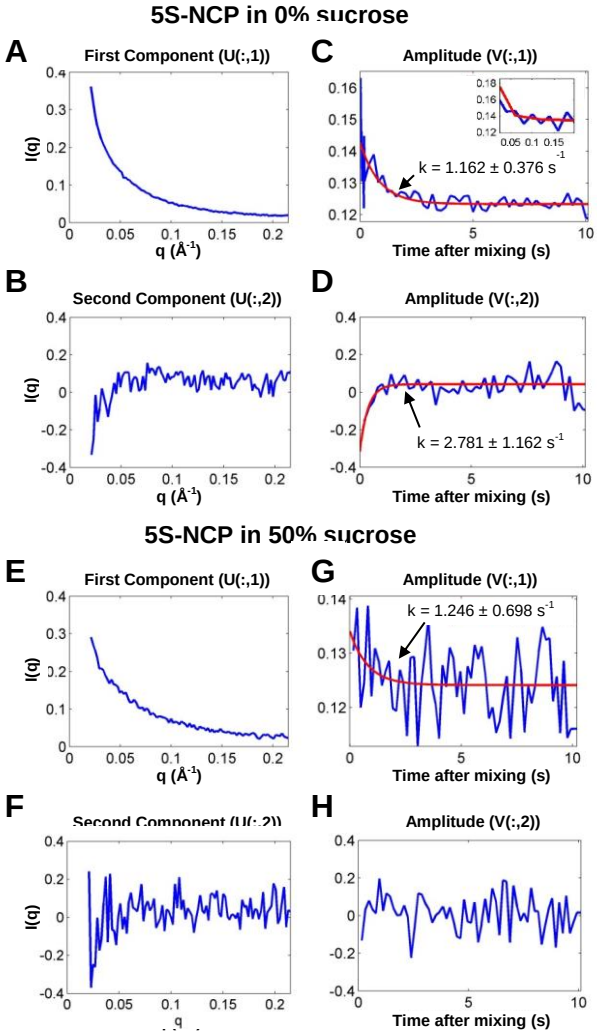
Supplementary Figure A.2 Comparison of P(R) curves calculated from models and equilibrium 601-NCP experiments. (**A**) Models with symmetrically released DNA from crystal structure 1AOI and theoretical P(R) curves for models shown. Four regions that provide physical insight into the conformational changes as the DNA is unwrapped are highlighted with colored boxes. The region in pink highlights the extension of the P(R) curves to the right that corresponds with the increasing largest dimension of the models. The region in blue reveals a formation of a peak at approximately 20 Å that corresponds with increasing single-helical extensions. The region in orange shows a decreasing peak at approximately 40 Å that reflects the decreasing overlap between the DNA ends (e.g. compare green to blue curves/models). The region in green shows a decreasing peak at approximately 80 Å that reflects the disruption of the overall wrapped structure. (**B**) Models with asymmetrically released DNA from crystal structure 1AOI and theoretical P(R) curves for models shown. (**C**) Experimental P(R) curves for 601-NCP in 50% sucrose and varying NaCl concentrations. Increasing the NaCl concentration induces conformational changes that reflect the same changes observed in the models.



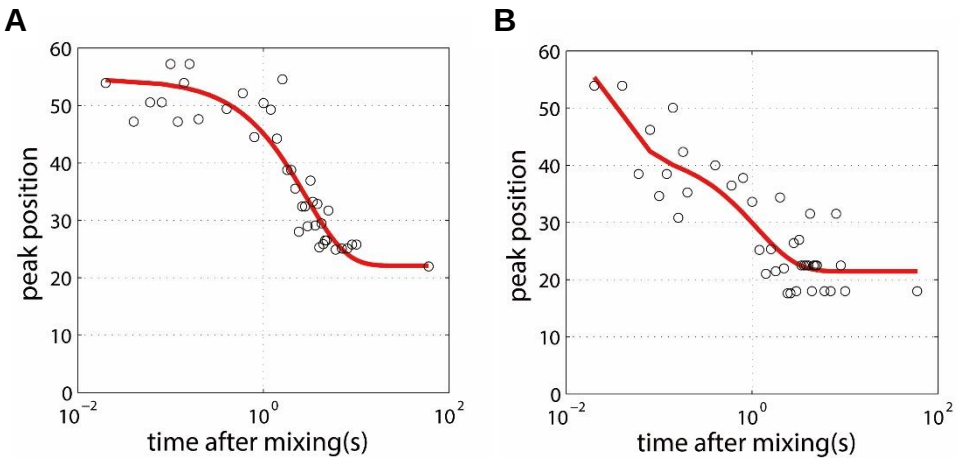
Supplementary Figure A.3 Time-resolved SAXS data for 601-NCP and 5S-NCP with and without sucrose**.** (**A,B**) Kratky plots ($I(q)xq^2$ vs. q) for NCPs dissociating in 1.9 M NaCl with 0% sucrose (protein and DNA visible) at the times indicated after mixing. For both (A) 601-NCP and (B) 5S-NCP, experimental data are shown as colored circles and regularized fits computed by GNOM are shown as solid black curves. Data has been offset to enhance visualization. (**C,D**) Kratky plots for NCPs dissociating in 1.9 M NaCl with 50% sucrose (only DNA visible) at the times indicated after mixing. For both (C) 601-NCP and (D) 5S-NCP, experimental data and fits are shown as colored circles and solid black lines, respectively, and the data has been offset to enhance visualization.



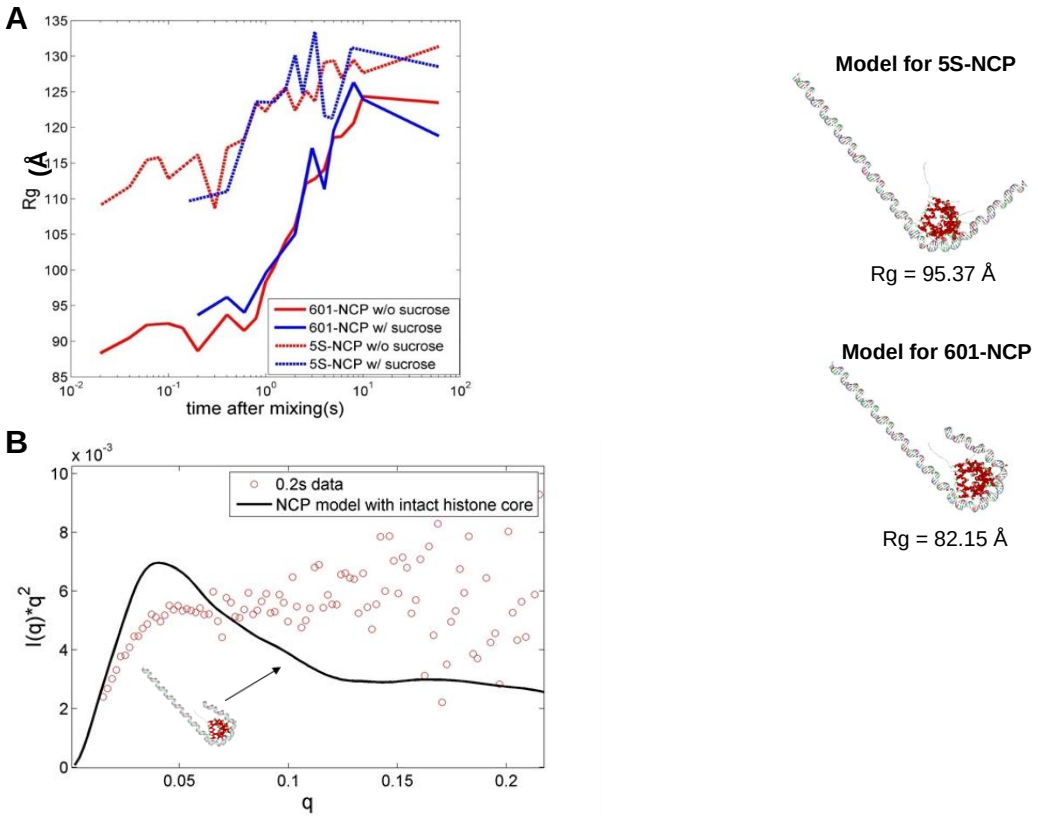
Supplementary Figure A.4 SVD analysis of TR-SAXS curves for 601-NCP in (**A-D**) 0% sucrose and (**E-H**) 50% sucrose. (**A,E**) The first basis components determined by SVD analysis of time binned SAXS curves. SVD analysis for time-resolved data was limited to regular linear profiles since TR-SAXS data has significantly more noise than static data. The first basis components dominate the other components by more than one order of magnitude. (**B,F**) The second basis components determined by SVD analysis of time binned SAXS curves. (**C,D,G,H**) Amplitudes of the components in (**A,B,E,F**) as a function of time after mixing. Because the first components dominate the SAXS profile changes, the amplitude changes (**C,G**) of the first components (**A,E**) correspond with the major transition (representing NCP disassembly). The time courses are well described by single exponential decay functions. Their nearly identical rates (k) between the sucrose and non-sucrose conditions indicate that sucrose has minimal effects on NCP dissociation dynamics. Analysis beyond the extraction of time constants from the major changes in the linear profiles was limited by data quality.



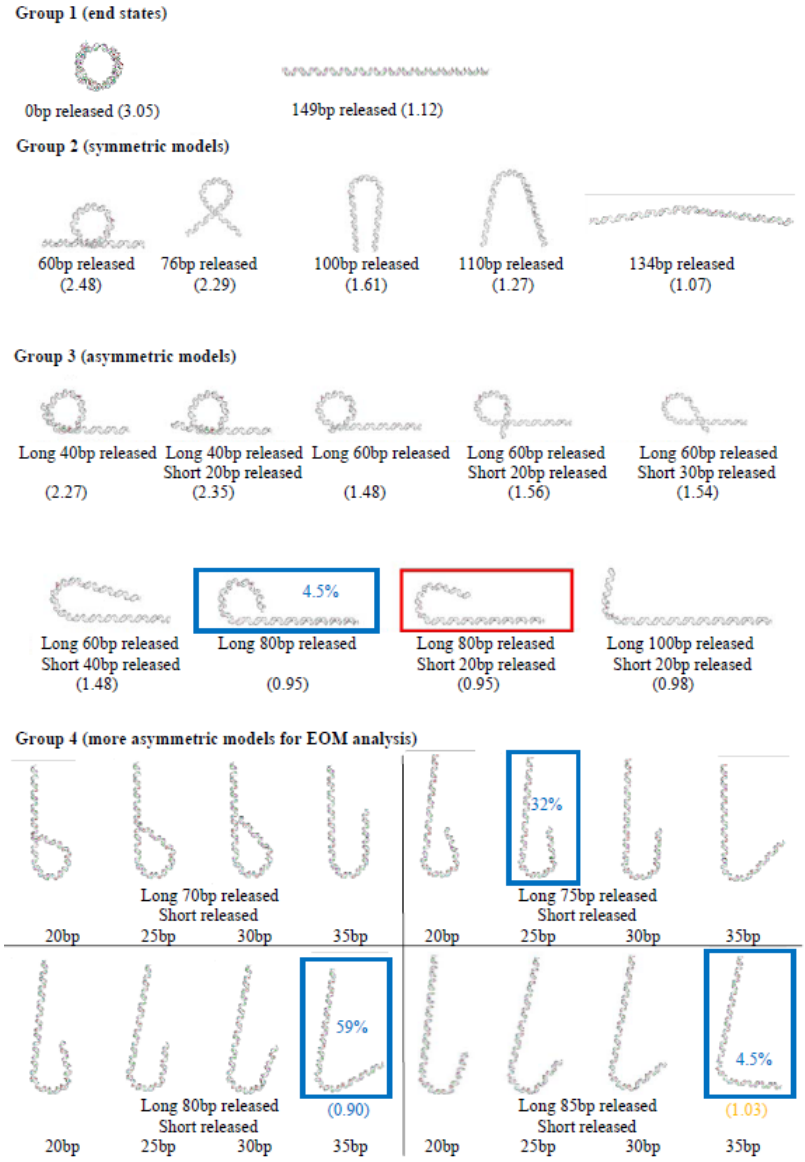
Supplementary Figure A.5 SVD analysis of TR-SAXS curves for 5S-NCP in (**A-D**) 0% sucrose and (**E-H**) 50% sucrose. (**A,E**) The first basis components determined by SVD analysis of time binned SAXS curves. SVD analysis for time-resolved data was limited to regular linear profiles since TR-SAXS data has significantly more noise than static data. The first basis components dominate the other components by more than one order of magnitude. (**B,F**) The second basis components determined by SVD analysis of time binned SAXS curves. (**C,D,G,H**) Amplitudes of the components in (**A,B,E,F**) as a function of time after mixing. Because the first components dominate the SAXS profile changes, the amplitude changes (**C,G**) of the first components (**A,E**) correspond with the major transition (representing NCP disassembly). The time courses are well described by single exponential decay functions. Their nearly identical rates (k) between the sucrose and non-sucrose conditions indicate that sucrose has minimal effects on NCP dissociation dynamics. Inset for (**C**): a fast phase of the amplitude change of the first component for 5S-NCP without sucrose. Analysis beyond the extraction of time constants from the major changes in the linear profiles was limited by data quality.



Supplementary Figure A.6 Major length scale analysis for (**A**) 601-NCP and (**B**) 5S-NCP in 0% sucrose**.** The position of the major peak in P(R) is plotted as a function of time. (**A**) 601-NCP, a single exponential fit yields a rate of 0.348 ± 0.082 s^{-1}. (**B**) 5S-NCP, a double exponential fit yields two rates: 36.4 ± 18.6 s^{-1} and 0.901 ± 0.354 s^{-1}. The fast phase observed for 5S-NCP without sucrose observed in SVD analysis (Supplementary Figure A.5C) agrees with that observed in the major length scale analysis (**B**) as well as I(0,t) analysis (Figure 3.5B).



Supplementary Figure A.7 Rg and Kratky evidence for a disrupted histone core. (**A**) Rg(t) of 601-NCP and 5S-NCP in 50% and 0% sucrose. For 601-NCP, the majority of the histone proteins dissociate after ≈ 1 s (Figure 3.5A). As a result, the DNA becomes the dominant scattering particle and Rg(t)s are quite similar between 0% and 50% sucrose. This was also observed for the 5S-NCP, except the majority of the histone proteins dissociate within the first ≈ 100 ms (Figure 3.5B). The 5S-NCP is generally more extended compared to the 601-NCP through the timescales observed. The Rgs for the 601-NCP 200 ms kinetic intermediate (Rg ≈ 90 Å) and 5S-NCP 160 ms ensemble average (Rg ≈ 112 Å) without sucrose are significantly larger than that computed for models of the NCP as shown in the two models on the right. The DNA structures in the two models were selected by single model chi square fits to the SAXS data and insight provided by P(R) analysis (Supplementary Figure A.8 and Figure 3.6). The models with the histone intact as an octamer on the wrapped end of the J-shaped DNA are insufficient for providing a sufficiently large Rg to match the Rgs observed. Assuming that all of the protein components remain attached (as evident in Figure 3.5A), the increased Rg may be explained by an extension of histone components away from the center of mass. This requires a disruption of the octameric histone and may be characterized by H2A-H2B dimers bound but positioned somewhere else along the DNA. (**B**) Comparison of Kratky profiles for the 601-NCP 200 ms kinetic intermediate in no sucrose (red circles) and the NCP model shown in the inset (black curve). The lack of a sharp peak in the data at q ≈ 0.045 $Å^{-1}$ indicates a structure without a compact globular core. The disagreement at high q is likely a result of scattering from the flexible disordered histone tails.



Supplementary Figure A.8 Chi square and EOM analysis of models to determine the conformation of the 200 ms 601-NCP kinetic intermediate and first 160 ms 5S-NCP**.** These symmetric and asymmetric DNA models were generated using Discovery Studio (Accelrys) and Nucleic Acid Builder. The curved portions of these models were migrated from the DNA component of crystal structure (pdb: 1AOI). $\overline{\chi^2}$ was calculated using Supplementary Eq. 5. The calculated $\overline{\chi^2}$s between the experimental SAXS profiles and the SAXS profiles calculated from the models are reported in parentheses below each model. Black $\overline{\chi^2}$s are calculated from fits to single models, whereas blue and yellow $\overline{\chi^2}$s are calculated from fits to the ensemble of models selected by EOM analysis. The best fitting single model using chi square analysis ($\overline{\chi^2} = 0.95$) is the 'J'-shaped asymmetric DNA (Long 80bp released, Short 20bp released) in the red box. The best fitting ensemble of models using chi square analysis ($\overline{\chi^2} = 0.90$) is highlighted by blue

boxes. The population of each model for the ensemble is reported as a percentage inside each box. The improvement in $\overline{\chi^2}$ value for the ensemble of models (0.90) compared to the single models (0.95) reveals better agreement with experimental data. Applying chi square analysis to the first 160 ms data for the 5S-NCP using single models and an ensemble of models selects the asymmetric DNA model (Long 85bp released, Short 35bp released) with $\overline{\chi^2} = 1.03$. This suggests that the 5S-DNA may also unwrap asymmetrically.

REFERENCES

[1] K. Luger, T.J. Rechsteiner, A.J. Flaus, M.M. Waye, T.J. Richmond, Characterization of nucleosome core particles containing histone proteins made in bacteria., J. Mol. Biol. 272 (1997) 301–311.

[2] L.M. Gloss, B.J. Placek, The effects of salts on the stability of the H2A-H2B histone dimer, Biochemistry. 41 (2002) 14951–14959.

[3] D.D. Banks, L.M. Gloss, Equilibrium Folding of the Core Histones: the H3-H4 Tetramer Is Less Stable than the H2A-H2B Dimer, Biochemistry. 42 (2003) 6827–6839.

[4] D.A. Hoch, J.J. Stratton, L.M. Gloss, Protein-Protein Förster Resonance Energy Transfer Analysis of Nucleosome Core Particles Containing H2A and H2A.Z, J. Mol. Biol. 371 (2007) 971–988.

[5] P.N. Dyer, R.S. Edayathumangalam, C.L. White, Y. Bao, S. Chakravarthy, U.M. Muthurajan, K. Luger, Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA, Methods Enzymol. 375 (2004) 23–44.

[6] D.I. Svergun, M.H.J. Koch, Small-angle scattering studies of biological macromolecules in solution, Reports Prog. Phys. 66 (2003) 1735–1782.

[7] D. Svergun, C. Barberato, M.H. Koch, CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 28 (1995) 768–773.

[8] P. Bernadó, E. Mylonas, M. V Petoukhov, M. Blackledge, D.I. Svergun, Structural characterization of flexible proteins using small-angle X-ray scattering., J. Am. Chem. Soc. 129 (2007) 5656–64.

Appendix B

Supporting Information for Chapter 4

Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core

B.1 Supplementary Methods

B.1.1 TR-SAXS Experiments

All TR-SAXS experiments were conducted at BioCAT Sector 18 at Advanced Photon Source (APS): x-ray energy =10 keV, sample-to-detector distance ≈ 2.5 m. The experimental procedures and SAXS image analysis are described in detail in [1]. Briefly, time resolution was achieved by incorporating a Biologic SFM-400 stopped flow mixer. Custom sample cells employed quartz capillaries (2 mm path length, 10 μm walls) from Hampton Research. NCPs in ≈ 0 M NaCl (none added) were mixed with high salt buffers (3.125 M NaCl) at a ratio of 2:3 (respectively) to achieve a final NaCl concentration of 1.9 M, and at a ratio of 3:2 to achieve a final NaCl concentration of ≈ 1.25 M. The final NCP concentration was ≈ 2-3 mg/mL (10-15 μM). Matching buffers were measured before and after each experiment by replacing the NCP sample syringe with 0 M NaCl buffer. Measurements were made with 0% and 50% sucrose.

A PILATUS 100K (Dectris) detector was operated in "movie mode" to capture the scattering patterns as a function of time. Since the samples showed radiation damage when exposed to the full x-ray beam at APS after 500 ms, two data collection schemes were employed to span short and long time points. For short timescales, images were

acquired using full beam for the first 500 ms in 5 ms intervals (2 ms exposure + 3 ms readout) for 0% sucrose, and 10 ms intervals (7 ms exposure + 3 ms readout) for 50% sucrose. For long timescales, 10 s of data was collected with an attenuated x-ray beam and an acquisition interval of 100 ms (95 ms exposure + 5 ms readout) for both 0% and 50% sucrose. Longer time scales (> 10 s) were limited due to backflow in the mixer.

B.1.2 TR-SAXS Data Analysis

All SAXS images were processed using MATLAB (MathWorks) as in [1]. For longer time points (> 50 ms), SAXS profiles from multiple frames were binned to improve the signal-to-noise ratio (at the expense of time resolution). SAXS profiles for corresponding time bins from subsequent repeats (4-10 for each condition) were averaged to improve the signal-to-noise ratio. I(0,t) was calculated using GNOM [2].

B.1.3 Generation of EOM DNA Models

The DNA pool included 9182 structural models generated using PyMol (expanded from 32 models in [1]). In addition to the fully wrapped (PDB: 1AOI) and fully extended (generated using Nucleic Acid Builder [3]) DNA conformations, 9180 DNA models were generated with varying degrees of unwrapping (with single base-pair resolution) by appending two linear DNA fragments to the two ends of one curved DNA component from 1AOI. Since each model consisted of three DNA components (two linear and one curved components) each had a minimum of 5 bp for accurate alignments. The linear DNA components contained 4 extra bp for alignment with the curved DNA component that were removed after stitching.

B.1.4 Minimal Ensemble Search (MES)

To independently validate the EOM-based result that multiple DNA species are required to fit the data, we ran Minimal Ensemble Search (MES) for the data acquired at 100 ms, 300 ms, and 300 s (1.2 M NaCl, 50% sucrose). MES identifies the minimal number of conformations required to fit the data [4]. For the 100 ms time point, the χ^2-value improved from 0.93 to 0.90 as a second DNA structure was added, however, further addition of structures beyond the second structure negligibly affected the χ^2-value. For the 300 ms time point, the χ^2-value improved from 0.99 to 0.96 as the number of structures increased from one to two. Additional structures did not improve the χ^2-value. The two structures selected by MES to best represent the data at 300 ms were the teardrop structure (91%) and the unwrapped structure (9%). The dominance of the teardrop structure at 300 ms agrees with the results of EOM. For the 300 s time point, the χ^2-value improved from 0.93 (one structure) to 0.90 (two structures). The two structures selected by MES included the teardrop (60%) and unwrapped (40%) structures. These results support the presence of the converging kinetic pathways to form the teardrop structure depicted in Figure 4.5B.

B.1.5 Kinetic Models for NCP Disassembly

The dynamic populations of NCP species observed by EOM at different salt conditions are summarized in the kinetic models shown in Figure 4.7 and Supplementary Figure B.10. In 1.2 M NaCl (Figure 4.7), the DNA rapidly unwraps from the histone octamer to form the "teardrop" conformation through two parallel pathways. In Pathway I, DNA unwraps from one end: Wrapped DNA Octasome (WO) → Asymmetric Open DNA

Octasome (AO) → Teardrop DNA Octasome (TO). In Pathway II, DNA unwraps from both ends and the more flexible end rewraps: WO → Symmetric Open DNA Octasome (SO) → TO. Instead of rewrapping, the SO can further unwind both DNA ends, and release dimers sequentially to form Open DNA Tetrasome (OT). However, since this transition (SO → OT) is ≈ 2.5 times slower than the rewrapping transition (SO → TO), the majority of NCPs (> 80%) form the TO after ≈ 0.5 s. The first H2A-H2B dimer dissociates in a step-wise manner through an open intermediate: TO → Teardrop DNA Open Intermediate (TI) → Teardrop DNA Hexasome (TH). Unwrapping of the TO → OT is a slow process (≈ 380 s) in 1.2 M NaCl. Although the hexasome forms within ≈ 30 seconds (TO → TH), the second dimer is protected from release by the teardrop DNA conformation for > 100 s (TH → OT). Tetrasome formation occurs after both DNA ends unwrap and is one order of magnitude slower than hexasome formation.

Parallel pathways are also observed in 1.9 M NaCl (Supplementary Figure B.10). However, at this more destabilizing condition, unwrapping of both ends (WO → SO) is the dominant pathway since it occurs nearly 10 times faster than unwrapping from only one end (WO → TO). Once both DNA ends unwrap, neither ends rewrap to form the teardrop conformation as observed in 1.2 M NaCl (note the directional difference of the transition between SO and TO in 1.2 M and 1.9 M NaCl). Consequently, TO is only formed through Pathway I and made up about 10% of the population in ≈ 0.6 s. NCPs with the teardrop DNA are short lived and the wrapped DNA end is quickly released after ≈ 1 s, resulting in the formation of more unwrapped octasomes (TO → SO). DNA unwrapping precedes dimer dissociation, and hexasomes and tetrasomes form sequentially after the DNA unwraps into open conformations (Symmetric Open DNA

Intermediate (SI) → SH → UT).

B.1.6 TR-FRET Experiments and Analysis

Two FRET pairs were employed: H3-78W-to-H2B-109Cys-AEDANS or H4-60W-to-H2A108Cys-AEDANS. This Trp to Cys-AEDANS FRET pair was chosen since the relatively short Forster radius of this dye pair (≈ 20 Å, compared to ≈ 50-60 Å for Cy3-Cy5) optimizes the system's sensitivity to dimer dissociation. Furthermore, the small size and small hydrophobic surface of the probes should reduce perturbations to the protein structure. Fluorophore incorporation was > 90%.

Time resolved data were collected using an AVIV ATF 105/305 spectrofluorometer interfaced with an AVIV stopped-flow tower. The excitation wavelength was 290 nm, and acceptor fluorescence was monitored at 500 nm. Dissociation reactions were initiated by diluting NCP solutions in 0 M NaCl into high salt buffers, either by stopped-flow (SF) or manual mixing (MM). SF data were collected from the dead time (5 ms) to 600 s at 1.2 M and 30 s at 1.9 M. To improve the signal-to-noise ratio, ≈ 20 kinetic traces were averaged for each SF experiment. The instrumentation set-up for MM allowed monitoring of the slower kinetic responses (from ≈ 2 s to 600-2000 s) with closure of the spectrometer chopper between time points to minimize photo-bleaching.

For NCP dissociation at NaCl concentrations of 1.2 and 1.9 M, datasets of 14 to 21 kinetic experiments were collected with varied final NCP concentrations of 10, 25, 50, 100, and 250 nM. The datasets for each FRET pair were fitted globally to a sum of three first-order exponentials using the Savuka 6.2.32 program [5]. In the global analyses, the relaxation times were linked across all kinetic experiments, thus treated as

independent of NCP concentration; the amplitudes associated with each phase were treated as local parameters for each kinetic trace.

The combination of manual mixing and stopped-flow mixing allowed kinetic responses to be monitored over time scales ranging from 10 ms to 1000 s (Supplementary Figure B.6). Initially, kinetic traces were fit individually to a sum of three exponentials, to assess the effect of NCP concentration on relaxation times and relative amplitudes. No significant dependence was observed for any parameter across this range of NCP concentrations. This independence is not unexpected given that the dissociation is a first order process, even though at 1 to 2 M NaCl, the overall stability of the NCP is concentration dependent (for example, [6]). Subsequent data analyses employed global fitting, which treated the relaxation times as independent of NCP concentration.

The kinetic parameters measured by FRET are summarized in Table 1. The errors (shown in parentheses) are reported at one standard deviation from these averaged values. For the relaxation times determined by FRET, the errors at one standard deviation were determined from a multi-dimensional analysis of the global fit of the data sets as described in ([5] pp. 37-54). The relative amplitudes are the average of values from individual traces, treated as local parameters in the global analyses of datasets containing 14 to 22 kinetic traces.

B.1.7 Interpretation of TR-FRET using two FRET pairs

Even though the two FRET pairs monitor different histone-histone interactions, the relaxation times determined by the two FRET pairs were quite comparable. The only

notable differences are the fastest and slowest relaxation times observed at 1.9 M. When the relaxation times were globally fitted, linked across all kinetic traces for both the H3-H2B and H4-H2A FRET pairs, the quality of the fits were similar to those for fits of the individual FRET pairs (reduced χ^2-values within 10%); the fitted relaxation times were 200 ms, 1.8 s and 8.3 s.

The fastest FRET relaxation time at 1.2 M NaCl, ≈ 2.5 s, is slower than the reactions which populate the partially unwrapped NCP structures detected by EOM analysis of the SAXS data (Figure 4.5). The relatively small amplitude of the fast phase (loss of ≈ 20% of the FRET signal) is consistent with a conformational change that opens the dimer-tetramer interfaces, rather than dissociation of an H2A-H2B dimer. Thus, there appears to be an uncoupling of DNA and histone conformational changes under mildly destabilizing NaCl concentrations. A similar "open" τ_{Open} is detected by FRET at 1.9 M NaCl. However, at this more destabilizing NaCl concentration, the protein-protein conformational changes are concomitant with the slower, large-scale DNA unwrapping reactions.

The slower relaxation times are associated with loss of ≈ 80% of the FRET signal. Surprisingly, the relative amplitudes for $\tau_{Hexasome}$ and $\tau_{Tetrasome}$ are unequal, despite both reactions corresponding to the loss of an H2A-H2B dimer. Furthermore, the two FRET pairs exhibit opposite patterns for whether the amplitude of $\tau_{Hexasome}$ is greater or less than the amplitude of $\tau_{Tetrasome}$. As elaborated below, comparisons of the amplitude patterns for the two FRET pairs are indicative of an asymmetric opening of the histone interfaces in the τ_{Open} phase (Supplementary Figure B.7). This asymmetry is consistent with the asymmetry of the DNA structures detected by EOM after ≈ 200 ms in 1.2 M

NaCl (Figure 4.5) and the asymmetric DNA unwrapping of the intermediates populated between 200 to 500 ms at 1.9 M NaCl (Figure 4.4).

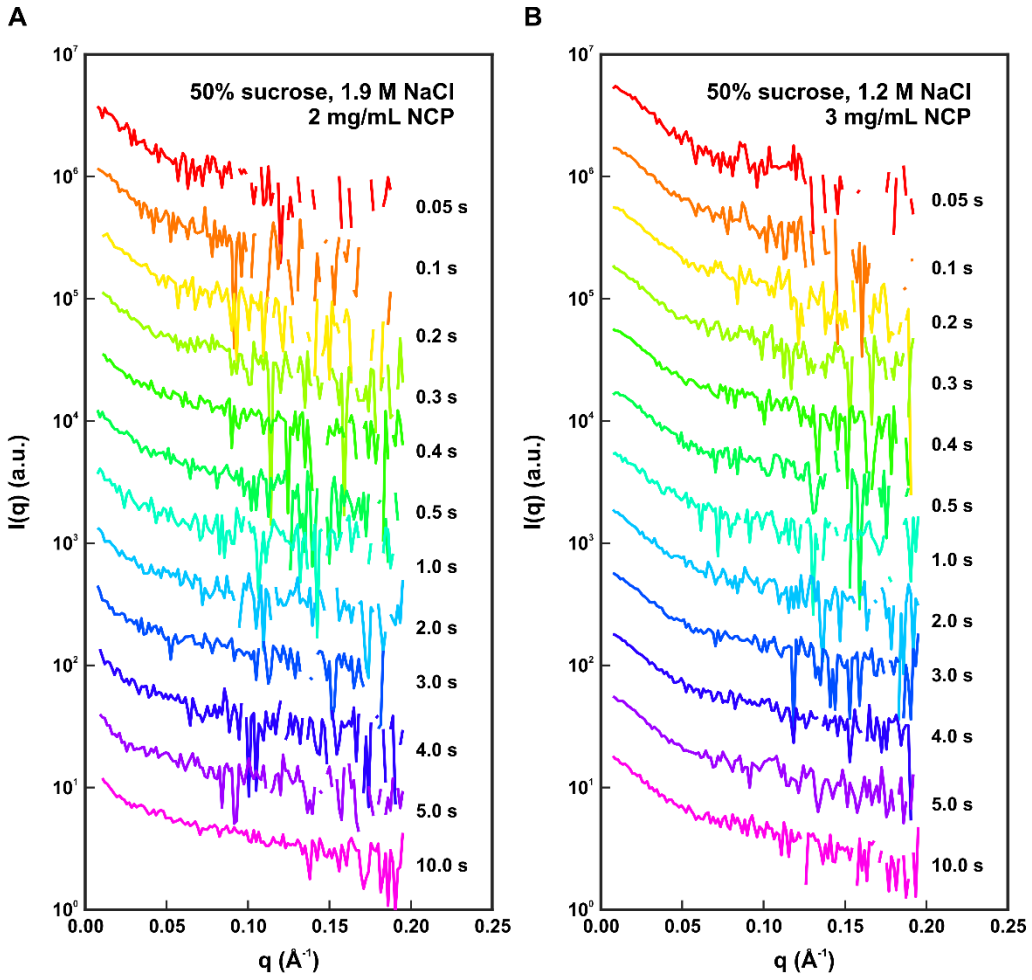
For the H3-H2B NCP, there are two symmetric donor-acceptor pairs (D-A and D'-A') which should contribute equally to the observed FRET signal (Figure 4.6A, Supplementary Figure B.7A). Thus, symmetric opening and subsequent formation of the hexasome and tetrasome would predict relative amplitudes of 20% for τ_{Open} (10% for D-A and 10% for D'-A') with 40% for both $\tau_{Hexasome}$ and $\tau_{Tetrasome}$ phases, arising from equal loss of FRET for dissociation of D-A and D'-A' dimers (illustrated in Supplementary Figure B.7D). The observed amplitude pattern of 20:30:50 can be explained by disruption of only one H2A-H2B/H3-H4 interface in an asymmetric τ_{Open} reaction, and then the affected dimer is released in the $\tau_{Hexasome}$ phase (Supplementary Figure B.7C). Given the position of the FRET pairs, the asymmetric opening is likely to affect the H2B-H4 four-helix bundle. Thus, dissociation of the first H2A-H2B dimer leads to a loss of 50% of the FRET in two steps (20% + 30%). Dissociation of the second dimer in the $\tau_{Tetrasome}$ reaction accounts for the other 50% decrease in the FRET amplitude (Supplementary Figure B.7C).

For the H4-H2A NCP, nearly equal changes in FRET signals are expected for the D-A and D-A' interactions (Supplementary Figure B.7B). A symmetric opening and dimer dissociation would predict similar FRET amplitudes for the $\tau_{Hexasome}$ and $\tau_{Tetrasome}$ phases (Supplementary Figure B.7F). The relative amplitude for the first dimer dissociation is significantly larger for than for the second, by a margin of 17 to 26% (Table 1). Despite an overall low amplitude, limited kinetic data for the H3-78W-to-H2A-108CA at 1.5 M NaCl showed a similar result of diminished amplitude for the

second dimer dissociation (relative amplitudes of 40:40:20 for three relaxation times). It is not possible to deconvolute the relative amplitudes into specific conformational changes for the same-face and cross-NCP FRET pairs. However, the results from the two H2A-CA containing FRET pairs suggests that increased mobility of the H2A tail contributes significantly to the τ_{Open} relaxation time and first dimer dissociation.

Even though the two FRET pairs (H3-H2B NCP and H4-H2A NCP) monitor different histone-histone interactions, the three relaxation times determined by the two FRET pairs were quite comparable (Table 1, Supplementary Figure B.7A-B). In contrast to the H3-H2B NCP, the relative amplitude for $\tau_{Hexasome}$ is significantly larger than that for the dissociation of the second dimer, $\tau_{Tetrasome}$. Although it is not possible to deconvolute the relative amplitudes into specific conformational changes for either FRET pair, the inequality of the amplitudes for the hexasome and tetrasome phases are consistent with the two dimers being released asymmetrically.

B.2 Supplementary Figures

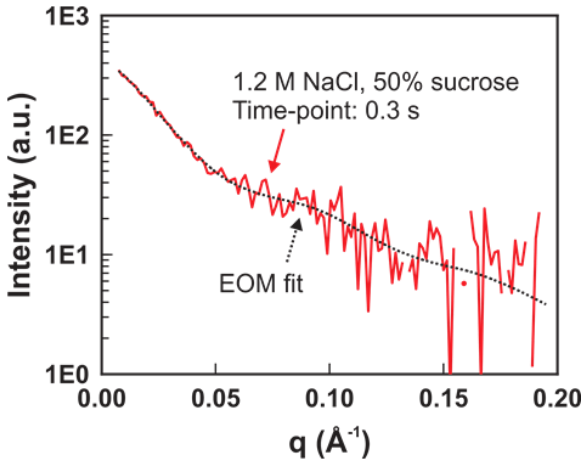


Supplementary Figure B.1 Time-resolved SAXS profiles for NCPs destabilized by NaCl in 50% sucrose. (**A-B**) Time-course measured for NCPs in 1.9 M NaCl (A) and 1.2 M NaCl (B). SAXS profiles for the time points after 0.5 s were collected using an attenuated x-ray beam to reduce the effects of radiation damage. Similar signal-to-noise ratios were achieved for the data collected within each salt series by using longer time bins (1 s) for the time points collected with the attenuator (for details, see Methods). SAXS profiles were offset for enhanced visualization and negative values are not shown on the log plot.

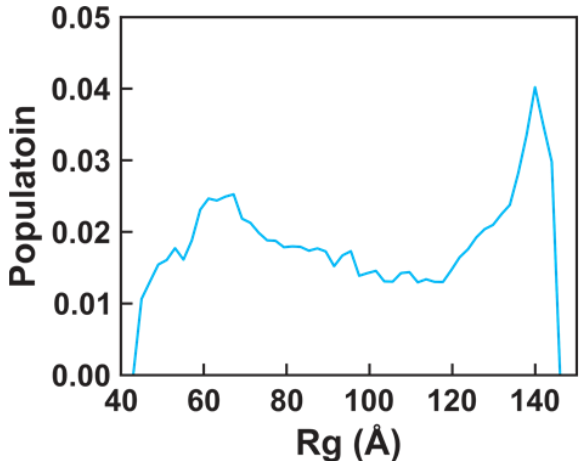
A

1.9 M NaCl		1.2 M NaCl	
Time (s)	χ^2	Time (s)	χ^2
0.05	1.12	0.01	1.05
0.1	0.99	0.02	1.05
0.2	0.99	0.04	0.94
0.3	0.89	0.06	0.86
0.4	1.09	0.08	1.05
0.5	1.02	0.1	0.99
1	1.03	0.2	0.94
2	1.1	0.3	1.03
5	1.28	0.4	0.86
10	1.9	0.5	1.13
		10	1.2
		600	1.2

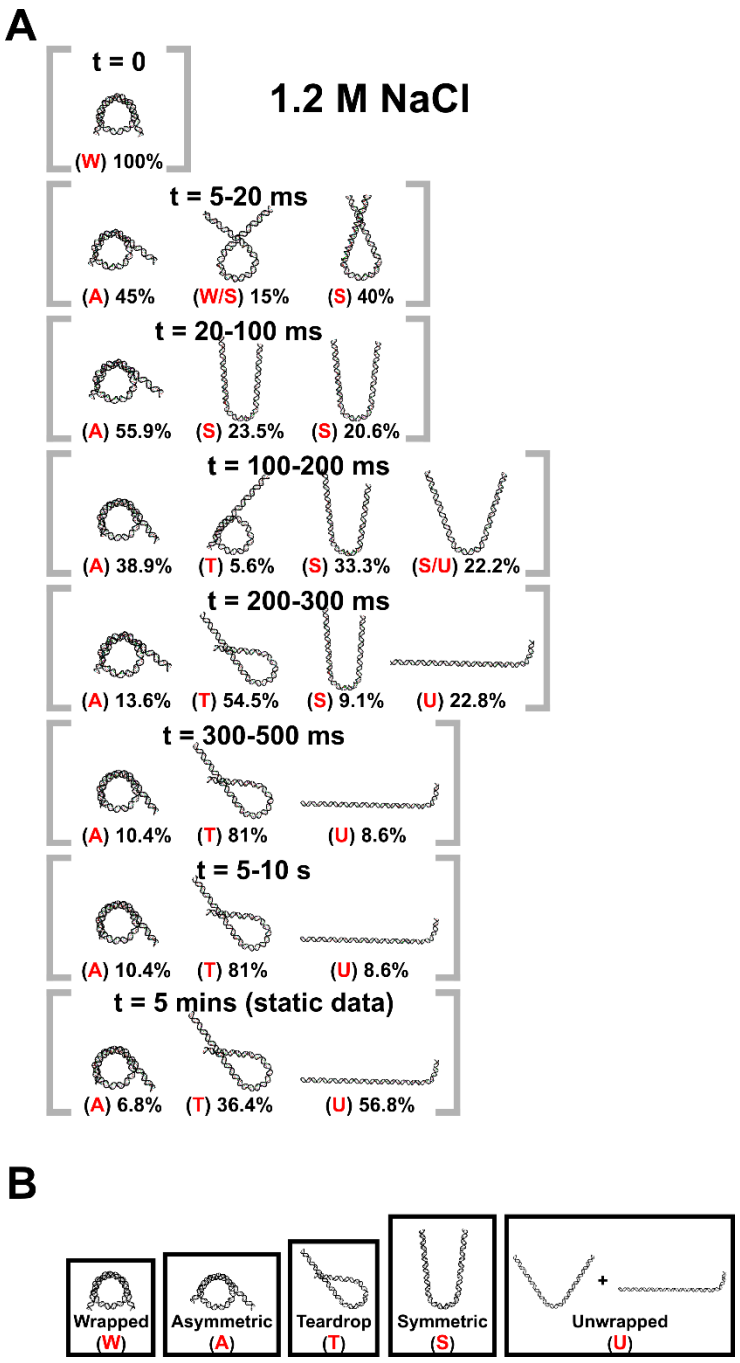
B



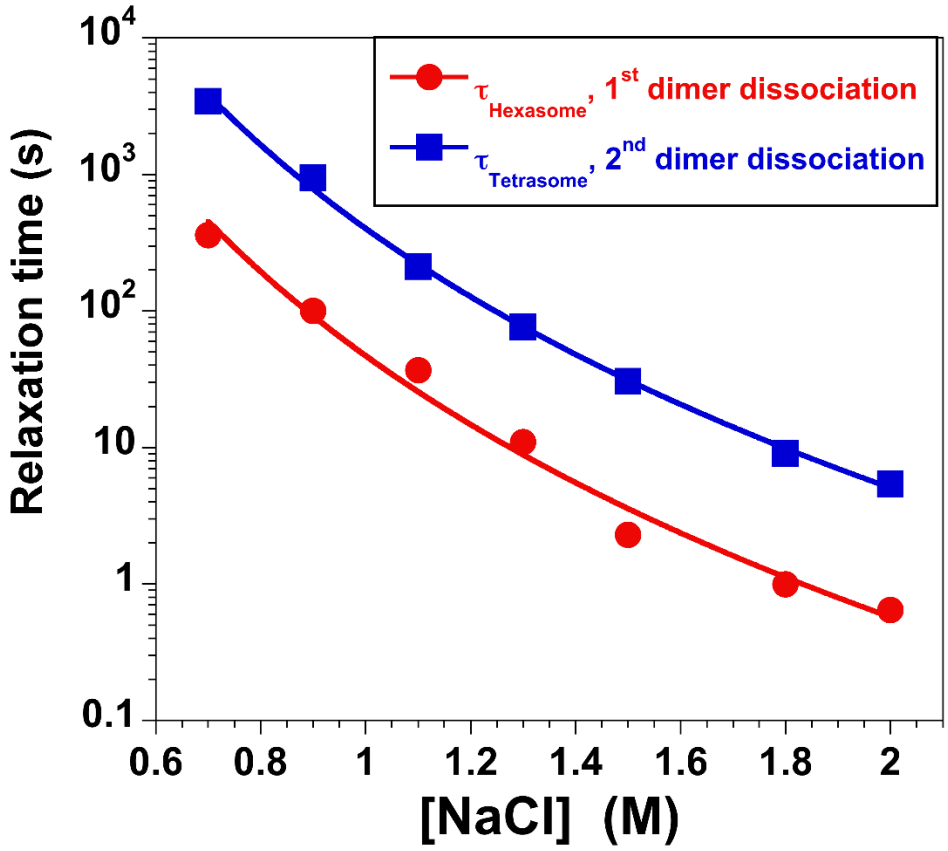
C



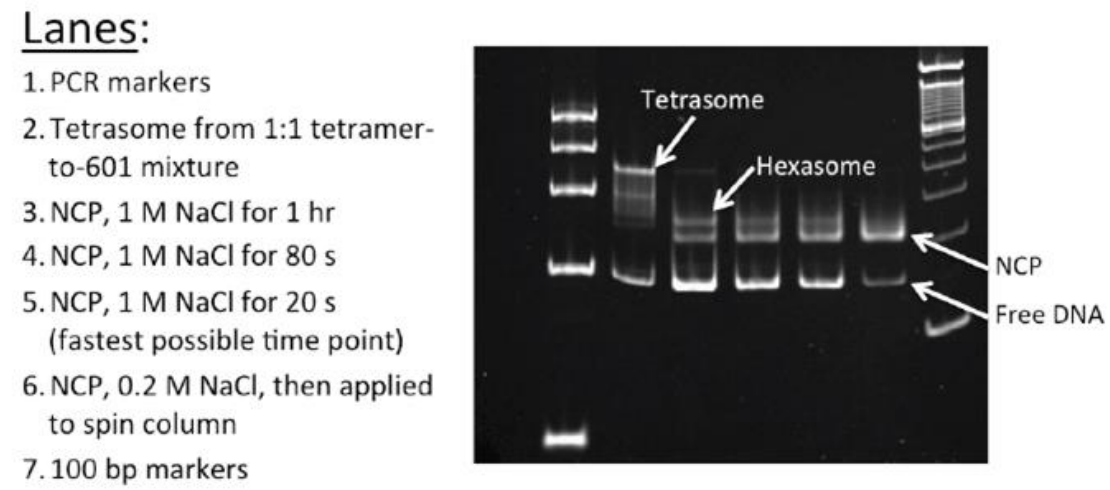
Supplementary Figure B.2 Ensemble optimization method (EOM) analysis of TR-SAXS data. (**A**) Table of χ^2-values for EOM fits to TR-SAXS data. (**B**) Representative fit of SAXS profile calculated for the ensemble of structures selected by EOM that best recapitulate the measured TR-SAXS profile. (**C**) Histogram of radius of gyration (Rg) values for the DNA pool (9,182 structures) used for EOM analysis. The Rg histograms for the time course of selected ensembles are shown in Figures 4.4A and 4.5A. The pool of DNA structures utilized in this study was sufficiently large enough to create ensembles of structures that fit the data well.



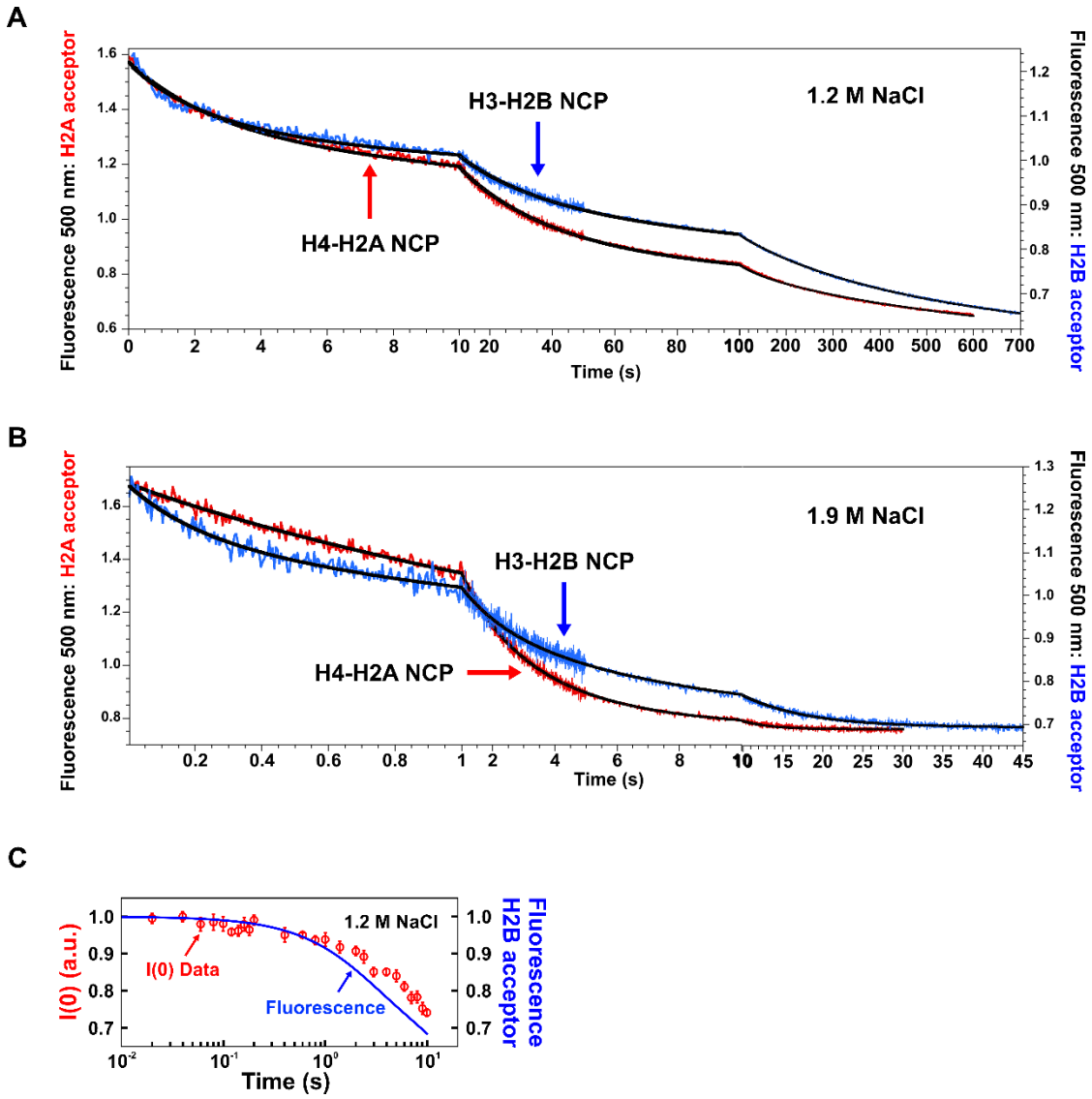
Supplementary Figure B.3 Ensembles of DNA structures selected by EOM. (**A**) DNA structures selected for NCPs in 1.2 M NaCl as a function of time. Percentages shown are the weights of the species out of the total population at the indicated time point. (**B**) Generalized classes of DNA conformations into which the models selected by EOM were grouped. The red letters for each model in (A) designate how the structures were grouped for subsequent analysis. Multiple letters mean that the population weight was split between the reported classes.



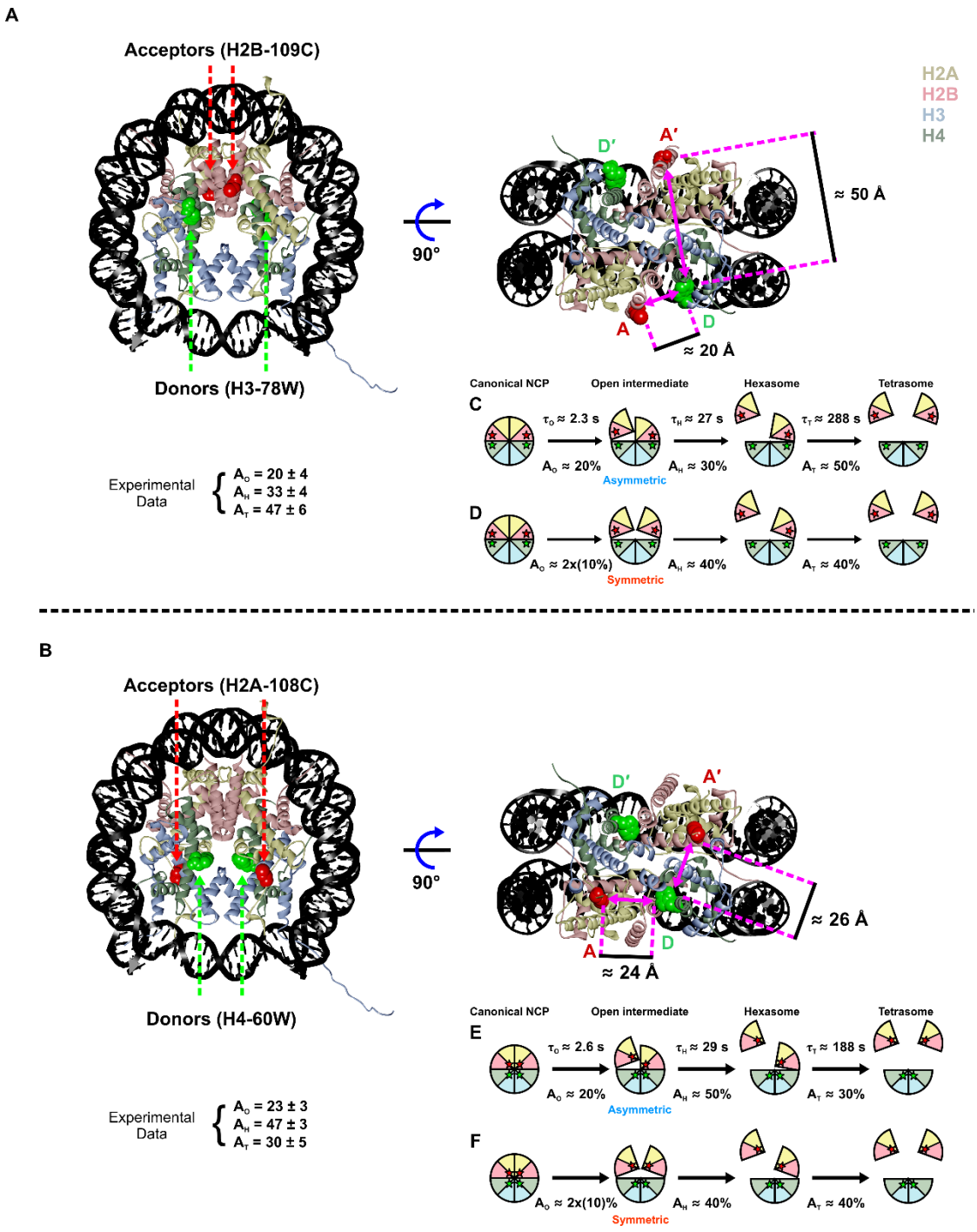
Supplementary Figure B.4 Salt-dependence of the kinetic responses measured by FRET using the H3-78W/H2B-109CA donor-acceptor pair. Dissociation was initiated primarily by manual mixing methods, with a time resolution of ≈ 2 s. However, at NaCl concentrations ≥ 1.5 M, the more rapid kinetic response was also determined using stopped-flow mixing, with a time resolution of ≈ 10 ms. Multiple kinetic traces at a given salt concentration were globally fit to a sum of two exponentials, yielding relaxation times $\tau_{Hexasome}$ (red circles) and $\tau_{Tetrasome}$ (blue squares) describing the dissociation of the first and second H2A-H2B dimers, respectively. The errors associated with the relaxation times are equal to or smaller than the size of the symbols. The lines are drawn to guide the eye and do not reflect a mechanistic fit of the data. Final conditions: 25 nM NCP, 20 mM Tris-Cl (pH 7.5), 0.1 mM EDTA, 0.1 mM DTT, 25 °C.



Supplementary Figure B.5 Gel assay for salt-induced dissociation of NCPs. Native gels for desalted NCPs incubated with NaCl for varying amounts of time show the timescales upon which hexasomes and tetrasomes are formed. Any amount of NaCl leads to the presence of a free DNA band on the gel (a technical artifact). These results qualitatively show that the hexasome is formed on the timescale of 20 to 80 s, with even slower formation of the tetrasome.



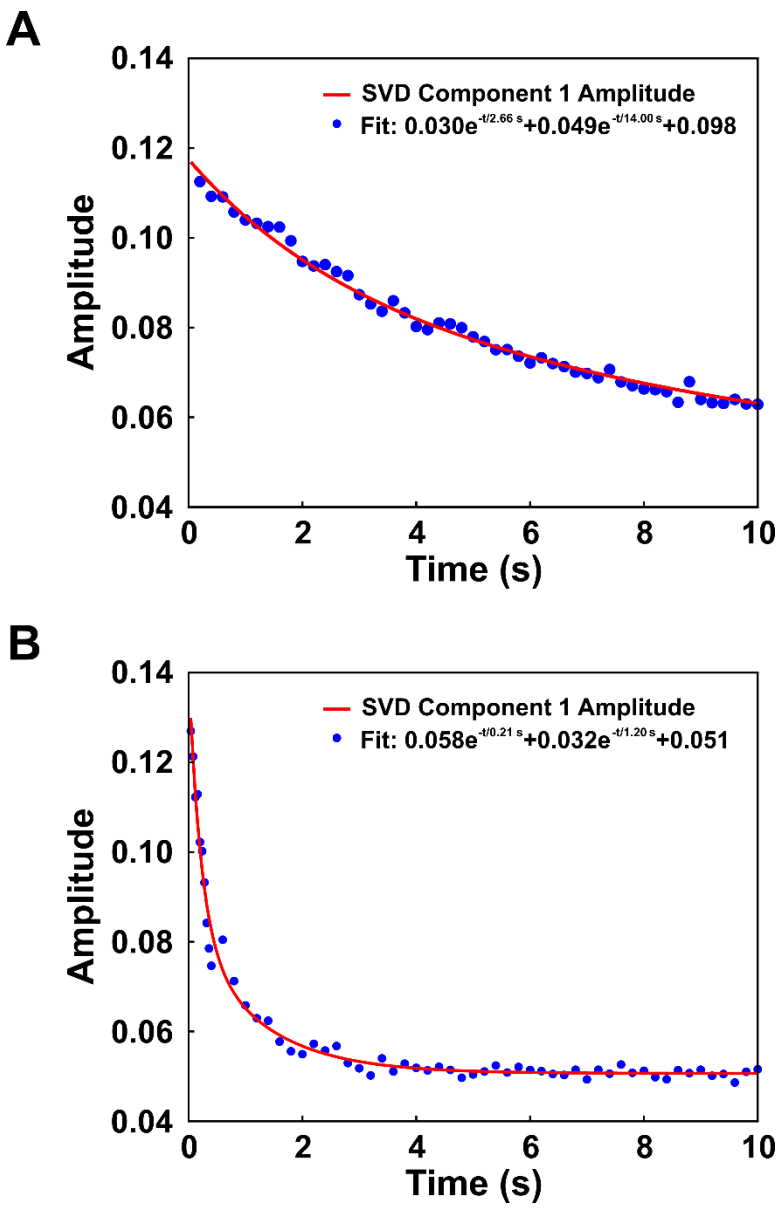
Supplementary Figure B.6 Representative FRET responses for NCPs destabilized by NaCl. (**A**) Acceptor fluorescence time-course measured for NCPs in 1.2 M NaCl. Each dataset (averaged over 20 kinetic measurements) shows the decrease in acceptor fluorescence after initiation of the salt jump for each FRET construct employed. The final NCP concentration was 250 nM NCP. (**B**) Same as (**A**), except NCPs were measured in 1.9 M NaCl. The solid black lines represent the global fits to a sum of three first-order exponentials for datasets collected as a function of NCP concentration (10 to 250 nM NCP). The relaxation times and relative amplitudes derived from these fits are reported in Table 1. Insets in (**A**) and (**B**) are log scale presentations of time-courses. (**C**) Comparison of relative TR-FRET changes with I(0,t) measured using TR-SAXS for NCPs in 0% sucrose and 1.2 M NaCl. The dynamics measured on overlapping time scales agree well between SAXS and FRET.



Supplementary Figure B.7 Models of dimer dissociation reported by two different FRET pairs. (**A**, **B**) NCP structure (1AOI) highlighting the distances between the FRET pairs: H3-78W donor to H2B-109CA acceptor (H3-H2B NCP) (**A**) and H4-80W donor to H2A-108CA acceptor (H4-H2A NCP) (**B**). Due to the symmetry of the structures in both constructs, distances are mirrored between the FRET pairs on both sides. Note: The Förster radius $R_0 \approx 20$ Å. The distances for the D-A pair (on the same NCP face)

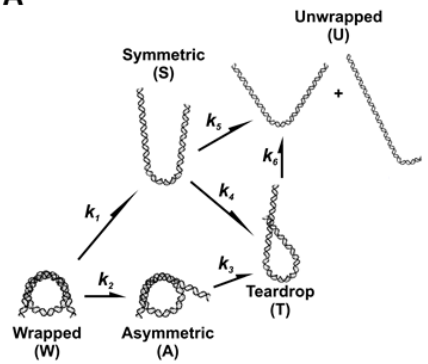
and D-A' pair (across the NCP) are similar for the H4-H2A NCP. Although equal changes in FRET signals are expected for the D-A and D-A' interactions, the dissociation of a single H2A-H2B dimer should still be reflected by the loss of ≈ 50% of the FRET signal due to the symmetry of the construct. Below each model in (**A**, **B**) are the measured FRET amplitudes (as reported in Table 1). (**C**-**F**) Models of possible histone configurations to explain the amplitude patterns observed for H3-H2B NCP (**C**-**D**) and H4-H2A NCP (**E**-**F**). For each FRET pair, the expected amplitude changes are shown for the scenarios that the open intermediate is formed asymmetrically (**C**, **E**) and symmetrically (**D**, **F**).

The measured amplitude patterns differ between H3-H2B NCP (33% → 47%) and H4-H2A NCP (47% → 30%). Release of the first H2A-H2B dimer may slightly destabilize the remaining H2A-H2B dimer (depicted as a slight separation) through the loss of some minor contacts between the two H2A-H2B dimers. The H4-H2A NCP construct may be more sensitive to this destabilization due to the positions of the FRET pair. Our data are most consistent with the sequential formation as follows: canonical NCP → asymmetric open intermediate → hexasome → tetrasome.



Supplementary Figure B.8 Determination of relaxation times using single-value decomposition (SVD) analysis of the full time-course of SAXS profiles. The time-dependent amplitude of the first component is shown for NCPs measured in 0% sucrose and 1.2 M NaCl (**A**) and 1.9 M NaCl (**B**). The amplitude dynamics of the first component represents the largest changes observed in the SAXS profiles. These changes correspond with NCP disassembly, and reflect both conformational changes as well as changes in molecular mass (changes in the forward scattering, I(0)). The decreasing amplitudes were fit to sums of exponential decays as shown. The relaxation times observed here are remarkably consistent with the values determined from FRET as shown in Table 1.

A



B

$$\frac{d[W]}{dt} = -k_1[W] - k_2[W]$$

$$\frac{d[S]}{dt} = k_1[W] - k_4[S] - k_5[S]$$

$$\frac{d[A]}{dt} = k_2[W] - k_3[A]$$

$$\frac{d[T]}{dt} = k_3[A] + k_4[S] - k_6[T]$$

$$\frac{d[U]}{dt} = k_5[S] + k_6[T]$$

C

time points (s)	W	S	A	T	U
0	1	0	0	0	0
0.02	0	0.4	0.45	0	0
0.1	0	0.44	0.56	0	0
0.2	0	0.33	0.39	0.06	0
0.3	0	0.09	0.14	0.55	0.23
0.5	0	0	0.1	0.81	0.09
10	0	0	0.025	0.825	0.15
600	0	0	0.07	0.36	0.57

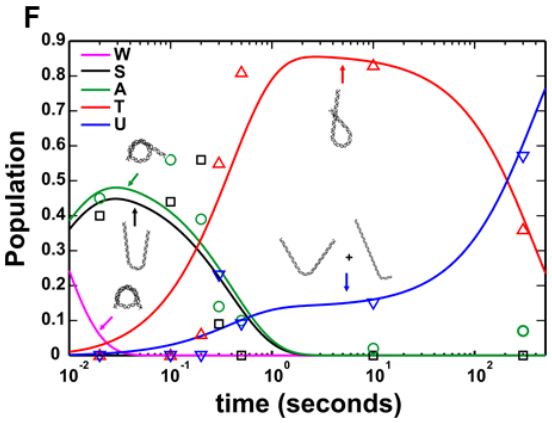
D

time points (s)	W or A	S	T	U
0	1	0	0	0
0.05	0.3	0.45	0	0.25
0.1	0.19	0.58	0	0.23
0.2	0.2	0.5	0	0.3
0.3	0	0.46	0.11	0.43
0.5	0.05	0.29	0.06	0.6
1.0	0	0.11	0	0.89
10	0	0	0	1

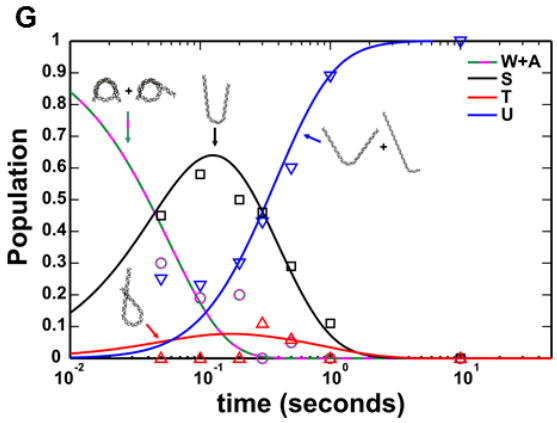
E

NaCl	k_1 (s^{-1})	k_2 (s^{-1})	k_3 (s^{-1})	k_4 (s^{-1})	k_5 (s^{-1})	k_6 (s^{-1})
1.2 M	68.4596	73.1314	2.4795	1.8739	0.7555	0.0026
1.9 M	15.8321	N/A	1.6067	-0.8103	2.8867	0.2401

F

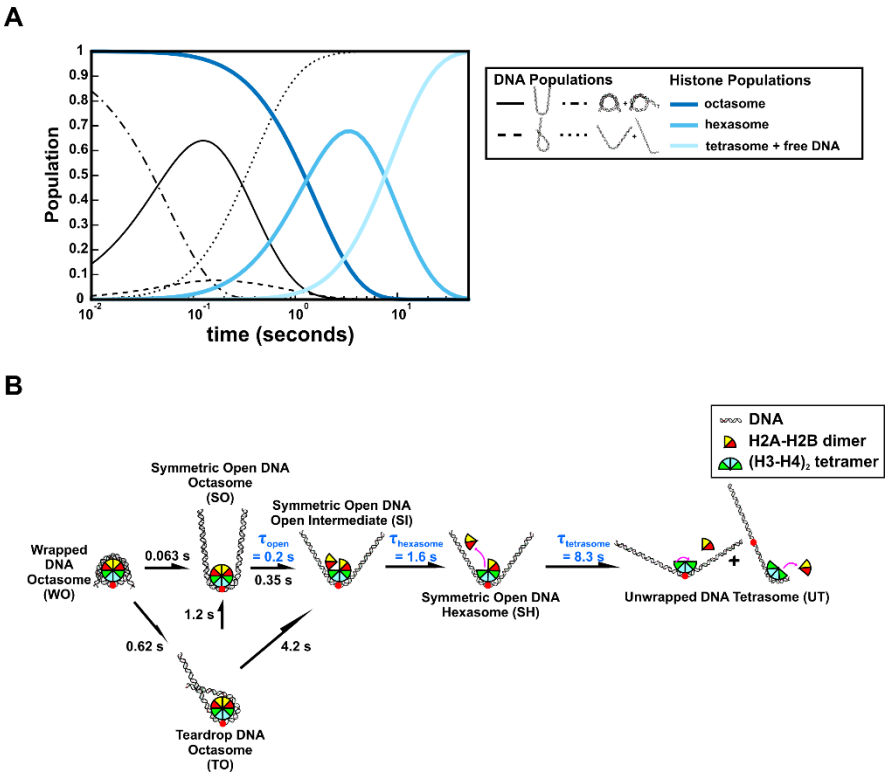


G



Supplementary Figure B.9 Calculation of kinetic rates from structures selected by EOM. (**A**) Kinetic scheme with pathways inferred from prominent DNA structures selected by EOM analysis for NCP data collected in 50% sucrose and 1.2 M NaCl (DNA structures selected at 1.9 M NaCl were nearly identical). (**B**) Kinetic equations describing the population dynamics for the kinetic scheme in (**A**). (**C**) Time-course of DNA structure populations determined from EOM analysis for NCP data collected with 50% sucrose and 1.2 M NaCl. (**D**) Same as (**C**), except using results for 1.9 M NaCl. (**E**) Rates for the kinetic scheme in (**A**) determined by solving the kinetic equations in (**B**) using population data for NCPs in 1.2 M (**C**) and 1.9 M (**D**) NaCl. Note: Convergent solutions were achieved using the same kinetic scheme for both datasets, suggesting that

the overall structures through which DNA unwraps are similar at both salt concentrations. However, the rates determined are quite different between these two conditions and indicates that not only is unwrapping faster at higher salt, but different pathways are preferred depending on [NaCl]. The opposite sign of k_4 at 1.9 M NaCl suggests that the symmetric unwrapped DNA is preferred over the teardrop at the more destabilizing condition. (**F**) Population data in (**C**) plotted with the simulated populations (lines) determined from the solved kinetic equations for 1.2 M NaCl (**E**). (**G**) Same as (**F**), except using results for 1.9 M NaCl (**D, E**). Note: the asymmetrically unwrapped DNA state (A) was not observed sufficiently enough in 1.9 M NaCl to be reliably fit, thus it was not included in the fitting and k_2 and k_3 are linked together.



Supplementary Figure B.10 Predicted populations and kinetic pathway of salt-induced NCP dissociation in 1.9 M NaCl. (**A**) Predicted populations of DNA conformational states (black lines) and NCP oligomerization states (blue lines) based on the kinetic rates determined from EOM analysis (Supplementary Figure B.9) and TR-FRET measurements (Table 1), respectively. (**B**) Kinetic models of salt-induced NCP dissociation in 1.9 M NaCl. The DNA conformations, associated pathways, and relaxation times (black numbers) were deduced from the SAXS EOM analysis (Figure 4.4). For each DNA conformation, the oligomeric state of the histone core was deduced from the lifetime of the hexasome and tetrasome populations (***A***). The relaxation times for dimer dissociation (blue numbers) were measured by TR-FRET (Table 1). For simplicity, histone orientations were centered on the dyad when possible. In 1.9 M NaCl, the majority of the DNA unwraps from both ends simultaneously, followed by dimer dissociation.

REFERENCES

[1] Y. Chen, J.M. Tokuda, T. Topping, J.L. Sutton, S.P. Meisburger, S. A. Pabit, L.M. Gloss, L. Pollack, Revealing transient structures of nucleosomes as DNA unwinds, Nucleic Acids Res. 42 (2014) 8767–8776.

[2] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.

[3] T.J. Macke, D.A. Case, Modeling Unusual Nucleic Acid Structures, in: N.B. Leontis, J. Santalucia (Eds.), Mol. Model. Nucleic Acids, American Chemical Society, Washington, DC, 1998: pp. 379–393.

[4] M. Pelikan, G.L. Hura, M. Hammel, Structure and flexibility within proteins as identified through small angle X-ray scattering, Gen. Physiol. Biophys. 28 (2009) 174–189.

[5] J.M. Beechem, Global Analysis of Biochemical and Biophysical Data, Methods Enzymol. 210 (1992) [2] 1-18.

[6] D.A. Hoch, J.J. Stratton, L.M. Gloss, Protein-Protein Förster Resonance Energy Transfer Analysis of Nucleosome Core Particles Containing H2A and H2A.Z, J. Mol. Biol. 371 (2007) 971–988.

Appendix C

Supplementary Information for Chapter 5

Nucleotide-dependent unwrapping of the nucleosome by the Chd1 chromatin remodeler

C.1 Methods

C.1.1 Generating DNA models for EOM

The success and reliability of ensemble optimization method (EOM) depends on the quality of the conformational pool from which the ensembles are selected. Ideally, models are informed by molecular dynamics or coarse-grained approaches that account for physical and biochemical data. In the case that this is not available, an alternate strategy, as done in the work described here, is to start from a known structure and manually modify the structure by replacing segments with altered conformations. Scripts that automate this process of splicing structural elements can quickly generate thousands of structures.

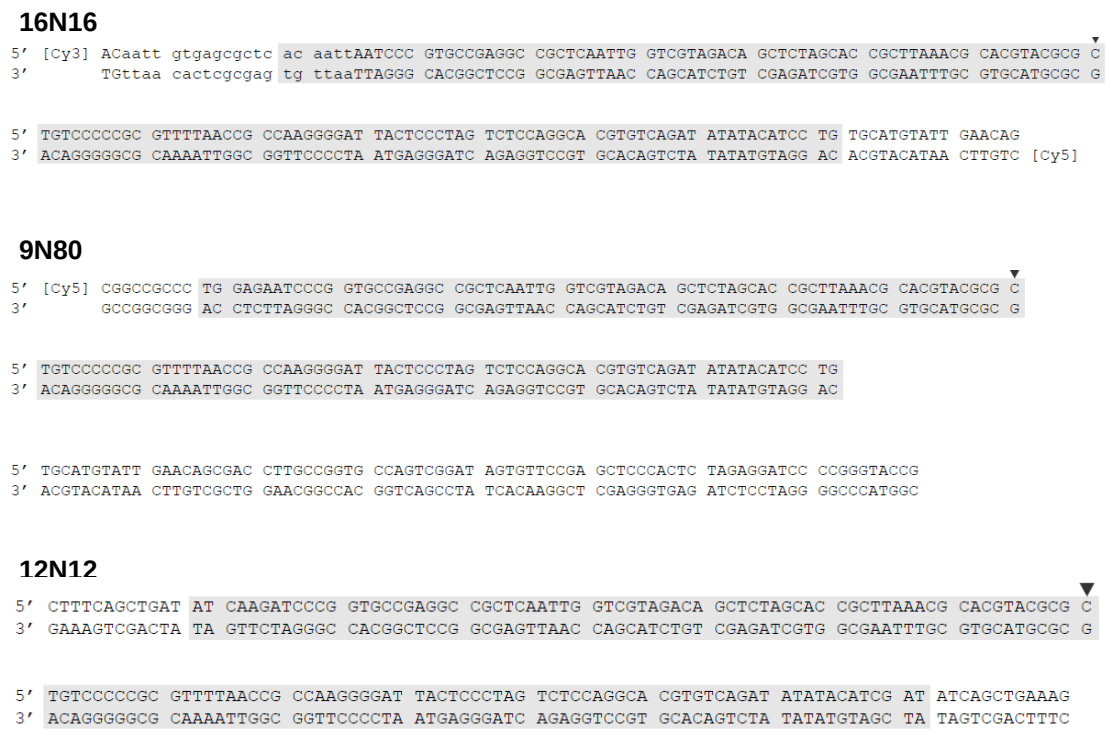
Fits to the data generally improves as more structures are considered. One difficult question to answer is, how many structures are enough? Since the ensemble search is essentially a mathematical exercise of minimizing discrepancies between the ensemble's theoretical profile and the experimental data, it is not obvious when "overfitting" occurs or when the quest for finding better mathematical solutions cause the selected ensembles to deviate from physiological relevance. Although there are no ways to resolve these concerns with absolute confidence, one can try running EOM with

different pools in different combinations to identify themes that persist between the structures selected from different pools. For determining a representative ensemble, it is advised to be conservative in terms of the number of models used to reduce the influence of outliers and the potential for overfitting. For drawing comparisons between ensembles selected using different SAXS datasets, it is safest to use the minimal library of structures that provides reasonable fits across all datasets.

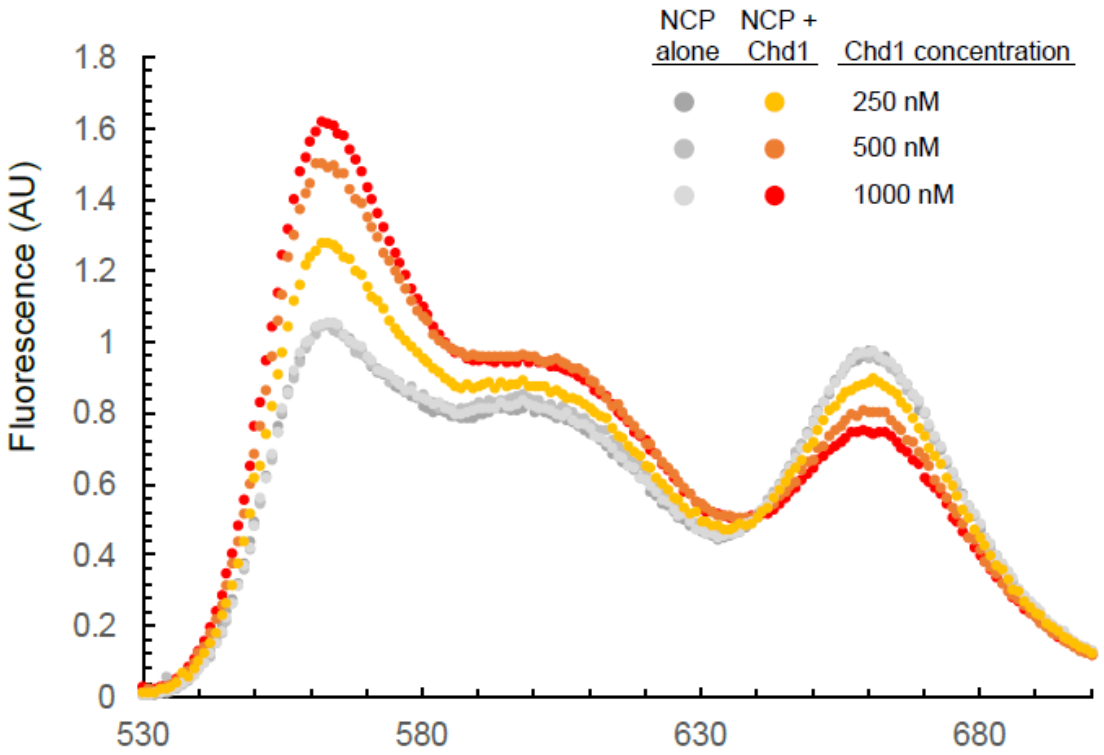
We initially tried using a DNA pool similar to the unwrapping DNA pool used in our previous work [1]. This pool was generated from the 147 bp NCP structure (1KX5) that was extended on both ends with an additional 11 bp for a total of 169 bp (12N12 nucleosomes). Scripts were written in Pymol to remove varying amount of DNA from the ends and append the same number of base pairs as linear B-form DNA. This DNA pool alone was insufficient to fit the data. Therefore, we generated a separate DNA pool that contained DNA kinks. These kinks allow the DNA to unwrap along trajectories other than that defined by the nucleosome.

A pool of kinked DNA was generated through the use of the 3D-DART online server [2]. The 12N12 nucleosome structure (extended 1KX5) was submitted to the server and a region of DNA was selected to be bent on either the left or right side of the dyad (Supplementary Figure C.5). The general strategy that motivated the choice of bending angle and regions of DNA to produce the kinked DNA pool is shown in Supplementary Figure C.6. Inclusion of the kinked DNA pool was found to be essential for obtaining good fits for the nucleosomes when Chd1 is present.

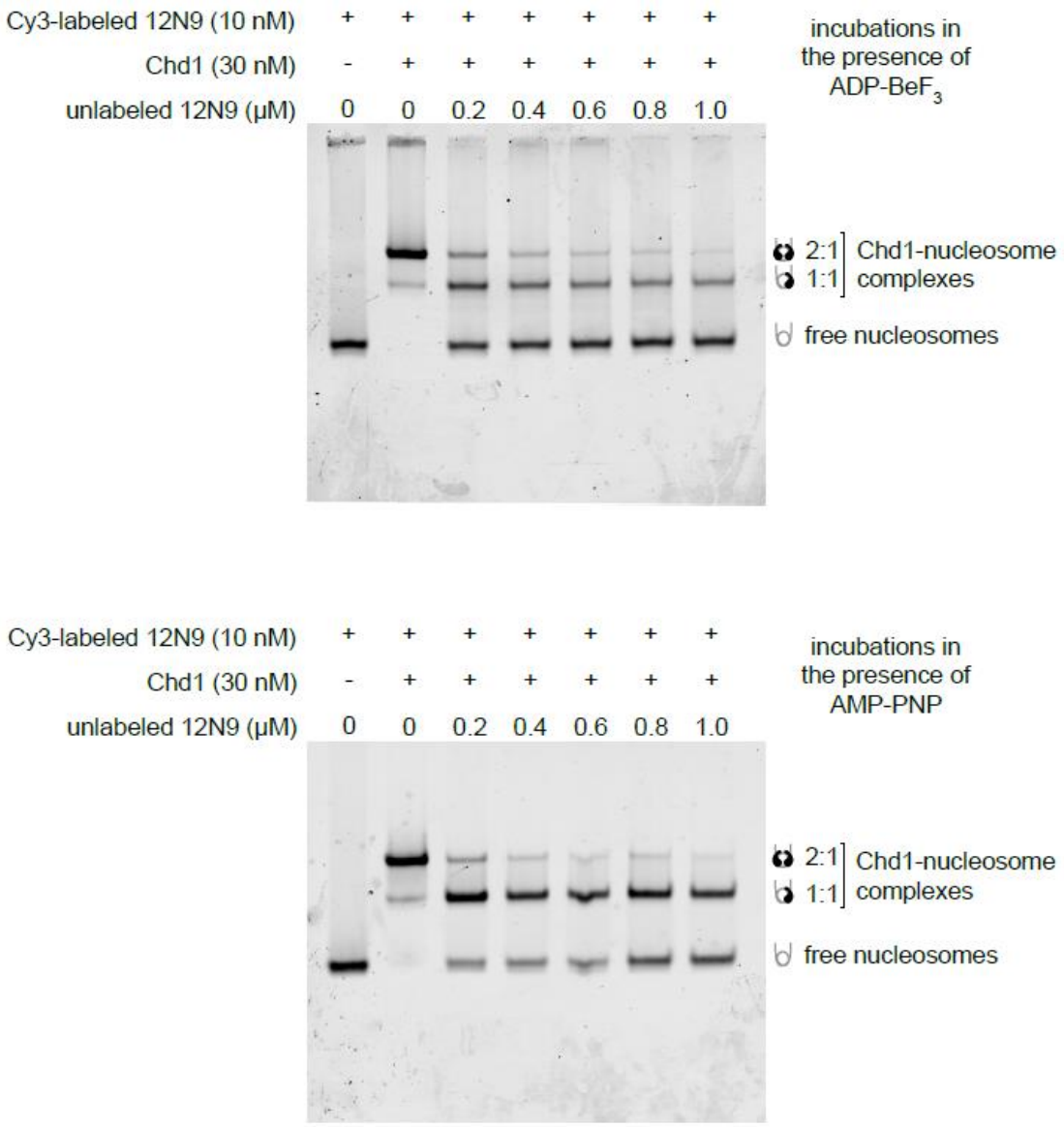
C.2 Supplementary Figures



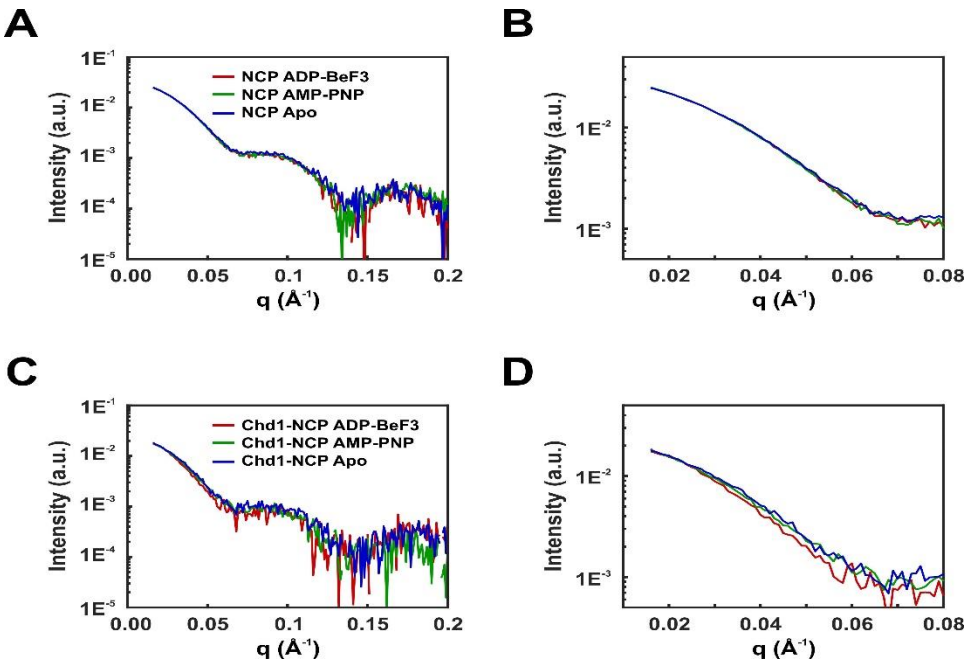
Supplementary Figure C.1 Nucleosomal DNA sequences used in this study. The core Widom 601 nucleosome positioning sequences (145 bp), which are wrapped around the histone core, are highlighted in gray. The dyad is indicated by a triangle. For the 16N16 sequence, lowercase letters indicate the location of the LacO site on one side.



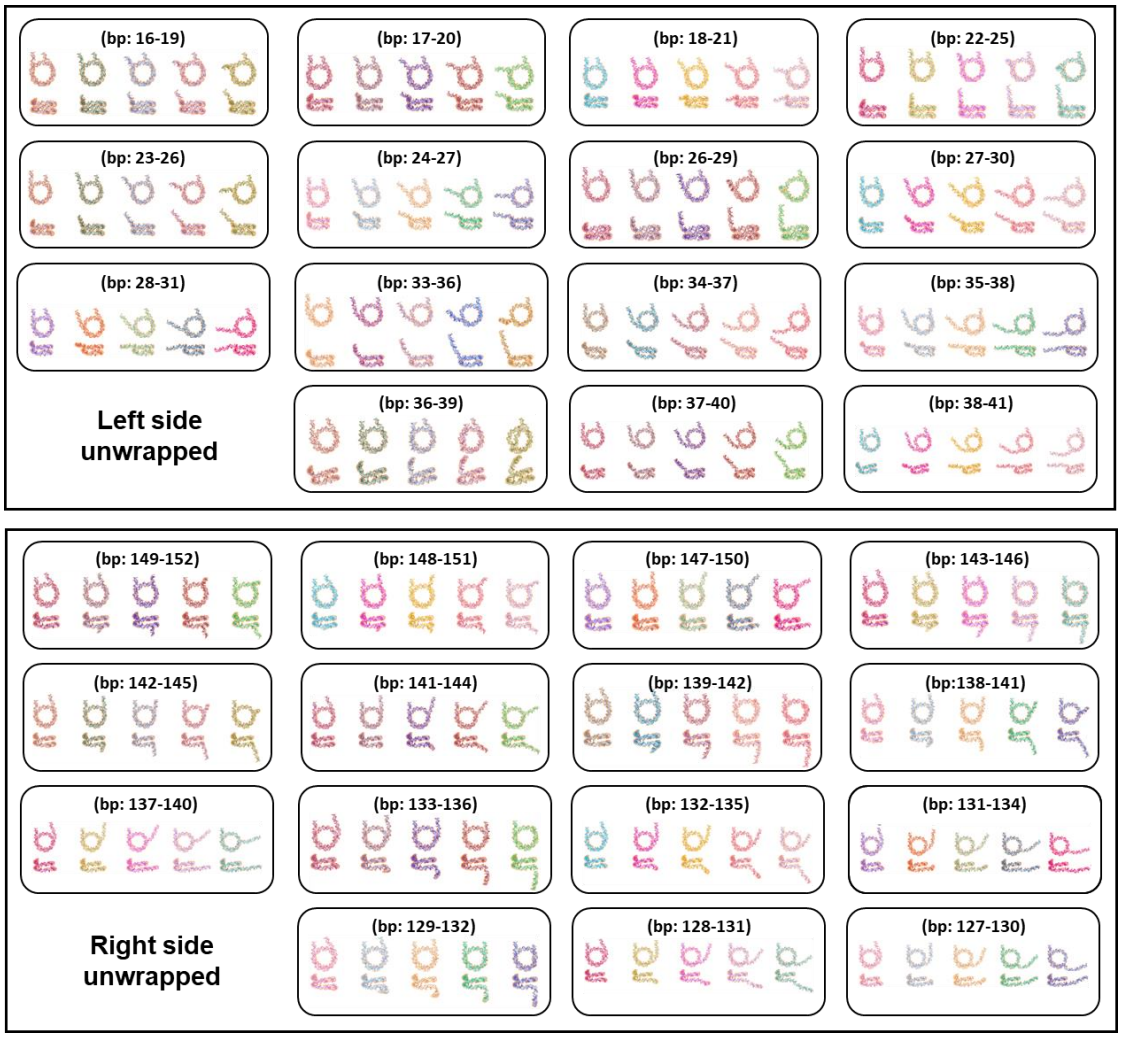
Supplementary Figure C.2 Chd1 unwraps 16N16 nucleosomes in the absence of nucleotide. Shown are wavelength scans of six independent samples containing either FRET-labeled 16N16 nucleosomes alone (gray) or 16N16 nucleosomes plus Chd1 in the absence of nucleotide (colored). In each case, addition of Chd1 yielded increased intensity of the Cy3 peak (560 nm) and decreased intensity of the Cy5 peak (660 nm), indicative of reduced FRET. Each pair of measurements were scaled to 560 nm intensity for nucleosome alone.



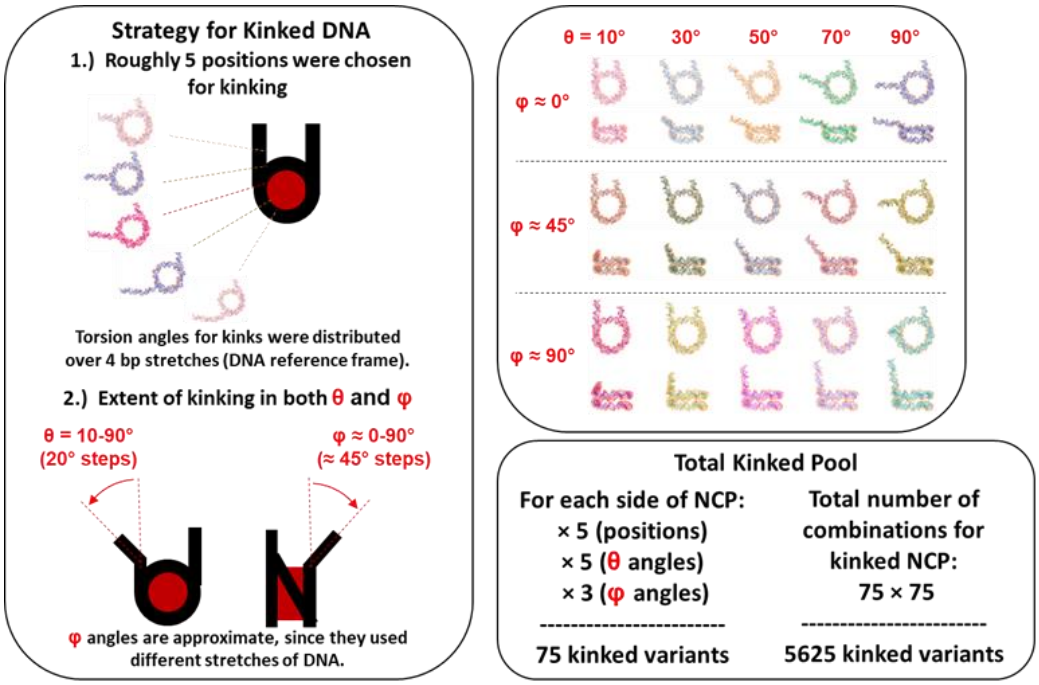
Supplementary Figure C.3 Chd1 binds preferentially to nucleosomes that are Cy3-labeled at H4(A15C). (**A**) Competition experiments in the presence of ADP-BeF_3, where Cy3-labeled 12N9 nucleosomes were incubated with Chd1 in the presence of increasing amounts of unlabeled 12N9 nucleosomes. Samples were separated by native PAGE and visualized on a Typhoon scanner. (**B**) Similar to (A), except incubation conditions contained AMP-PNP. The gels shown in (A) and (B) are representative of three independent measurements.



Supplementary Figure C.4 Comparison of SAXS profiles measured in the presence of different nucleotides and 60% sucrose**.** (**A**) SAXS profiles for NCP (12N12) in ADP-BeF_3, AMP-PNP, and Apo states. (**B**) Closer look at low-q region in (**A**). (**C**) SAXS profiles for Chd1-NCP (12N12) in ADP-BeF_3, AMP-PNP, and Apo states. (**D**) Closer look at low-q region in (**C**). NCP conformations remain unchanged with respect to different nucleotide states, but Chd1-NCP conformations show slight variations.



Supplementary Figure C.5 Variations of DNA bends used to generate kinked DNA pool with both in-plane and out-of-plane unwrapping. DNA bends were introduced into 12N12 nucleosomes using the 3D-DART server [2] on either the left or right side of the nucleosome. Bending angles of 10°, 30°, 50°, 70°, or 90° were applied as depicted in the 5 structures shown in each box. Each of these bends were applied across 4 basepairs as indicated. Note: some basepair ranges lead to varying degrees of out-of-plane unwrapping. Each variation of left- and right-side unwrapping was combined (by stitching left and right halves of structures together at the dyad) to produce 5,625 total variations of kinked DNA structures



Supplementary Figure C.6 Two angles were considered for the bending: θ, which represented the in-plane bending angle; and φ, the out-of-plane bending angle. θ ranged from 10-90° (20° steps) and φ ranged 0-90° (45°steps). The resulting variations for the left and right sides of the NCP are shown in Supplementary Figure C.5. The left and right halves of the NCPs were combined in every combination for a total of 5,625 kinked DNA variants.

REFERENCES

[1] Y. Chen, J.M. Tokuda, T. Topping, S.P. Meisburger, S.A. Pabit, L.M. Gloss, L. Pollack, Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core, Proc. Natl. Acad. Sci. 114 (2017) 334–339.

[2] M. van Dijk, A.M.J.J. Bonvin, 3D-DART: A DNA structure modelling server, Nucleic Acids Res. 37 (2009) 235–239.